DOCUMENT RESUME

ED 308 073 SE 050 656

AUTHOR Griffiths, Alan Keith
TITLE An Investigation of the Hierarchical and Developmental Structure of Selected Science Concepts.
SPONS AGENCY Social Sciences and Humanities Research Council of Canada, Ottawa (Ontario).
PUB DATE Apr 87
GRANT 410-80-0400
NOTE 102p.; Drawings or figures with small print may not reproduce well.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS Biology; Chemistry; *Cognitive Development; Conservation (Concept); Foreign Countries; Geology; High Schools; Physics; Piagetian Theory; Science Curriculum; Science Education; *Scientific Concepts; *Secondary School Science; *Vertical Organization
IDENTIFIERS Canada

ABSTRACT

This report describes the results of an attempt to identify a learning hierarchy for each of a number of science concepts mainly encountered first in the high school grades. The concepts studied relate to stoichiometric calculations and molarity, both from chemistry; to food web relationships and problems involving Mendel's laws from biology; to problems involving conservation of mechanical energy from physics, and to rock classification problems from geology. A hierarchy validated in terms of its inclusion and transfer characteristics is reported in each case except the last. Further, each hierarchy was examined for the relationship between learner developmental level and acquisition of its component skills. The conformation of the data was not suitable for such examination for two of the hierarchies, namely the conservation of mechanical energy hierarchy and the Mendel's laws hierarchy. For each of the other three hierarchies it was found that formal thought was not necessary to the acquisition of skills which, from a developmental perspective, would have been predicted to be dependent upon it. For each hierarchy data were collected from between 150 and 200 subjects. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED308073

An Investigation of the Hierarchical and Developmental Structure of Selected Science Concepts

by

Alan Keith Griffiths

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Alan K Griffiths

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A report of a study completed through research grant 410-80-0400 awarded by the Social Science and Humanities Research Council of Canada

Memorial University of Newfoundland
St. John's, Newfoundland, Canada A1B 3X8

April, 1987

SE 050 656

BEST COPY AVAILABLE

TABLE OF CONTENTS

ABSTRACT		iv
ACKNOWLEDGEMENTS		v
LIST OF TABLES		vi
LIST OF FIGURES		viii
CHAPTER		PAGE
1.	Theoretical background	1
2.	EXPERIMENTAL METHODS-RELATED LITERATURE	8
	Hierarchy Validation-Early Methods	8
	Hierarchy Validation-Recent Inclusion Methods	9
	Hierarchy Validation-Transfer Methods	12
	Developmental Level-Group Testing	19
3.	PROCEDURES	21
	Design	21
	Hierarchy Development	23
	Hierarchy Test Development	24
	Remedial Unit Development	24
	Developmental Test Development	25
	Statistical Tests	29
4.	HIERARCHY VALIDATION	30
	The Stoichiometry Hierarchy	30
	The Molarity Hierarchy	40
	The Conservation of Mechanical Energy Hierarchy	51
	The Food Web Hierarchy	63
	The Mendel's Laws Hierarchy	66

5. REQUIREMENT FOR FORMAL THOUGHT 76

The Stoichiometry Hierarchy 76

The Molarity Hierarchy 79

The Food Web Hierarchy 82

The Conservation of Mechanical Energy and Mendel's laws Hierarchies 85

6. EPILOGUE AND PROLOGUE 86

ABSTRACT

This report describes the results of an attempt to identify a learning hierarchy for each of a number of science concepts mainly encountered first in the high school grades. The concepts studied relate to stoichiometric calculations and molarity, both from chemistry; to food web relationships and problems involving Mendel's laws from biology; to problems involving conservation of mechanical energy from physics, and to rock classification problems from geology. A hierarchy validated in terms of its inclusion and transfer characteristics is reported in each case except the last. Further, each hierarchy was examined for the relationship between learner developmental level and acquisition of its component skills. The conformation of the data was not suitable for such examination for two of the hierarchies, namely the conservation of mechanical energy hierarchy and the Mendel's laws hierarchy. For each of the other three hierarchies it was found that formal thought was not necessary to the acquisition of skills which, from a developmental perspective, would have been predicted to be dependent upon it. For each hierarchy data were collected from between 150 and 200 subjects.

ACKNOWLEDGEMENTS

The research described in this report would not have been possible without the assistance of many people. The author is deeply grateful to former graduate students Bette Grant, Barbara Hopkins, John Pottle, Bill Roache and Pat Whelan; to research assistants Dr. Roda Amaria and Kevin Thomey, to faculty colleagues Dr. Glen Clark, Dr. Michael Collins, Dr. Colin Davies and Harvey Weir; to the Avalon Consolidated School Board, the Roman Catholic School Board for St. John's and the Pentecostal Assemblies School Board, tens of their teachers and thousands of their pupils; and to Mrs. Winnifred Martin for her patience and skill in typing the manuscript. Finally, the generosity of the Social Sciences and Humanities Research Council of Canada in providing the research funds without which this work could not have taken place, is gratefully acknowledged.

LIST OF TABLES

TABLE		PAGE
1	Comparison of Group and Individual Testing	28
2	Hierarchy Sub-tests	33
3	Ordering-Theoretic Method: Percentage of Exceptions to Hierarchical Connections, Stoichiometry Pretest	34
4	Dayton and Macready Analysis for Stoichiometry Sub-Hierarchies	36
5	Transfer Test, Stoichiometry Hierarchy	42
6	Tetrachoric Correlation Coefficient $r_{(tet)}$	45
7	Ordering-Theoretic Method: Percentage of Exceptions to Hierarchical Connections, Molarity Pretest	47
8	Dayton and Macready Analysis for Molarity Sub-Hierarchies	49
9	Test of Transfer Molarity Hierarchy	50
10	Phi Coefficients Between Items Testing the Same Skill on the Pretest, Mechanical Energy Hierarchy	55
11	Ordering-Theoretic Method: Percentage of Exceptions to Hierearchical Connections Conservation of Mechanical Energy Pretest	56
12	Dayton and Macready Analysis for Kinetic-Potential Energy Sub-Hierarchies	58
13	Test of Transfer for Conservation of Mechanical Eneregy Hierarchy	61
14	Tetrachoric Correlation Coefficients Pre- and Posttests Food Web Hierarchy	64
15	Ordering-Theoretic Method: Percentage of Exceptions, Pre- and Posttest Food Web Hierarchy	67

16	Test of Transfer for Food Web Hierarchy	68
17	Ordering-Theoretic Method: Percentage of Exceptions to Hierarchical Connections, Mendel's Laws Hierarchy	70
18	Dayton and Macready Analysis for the Mendel's Laws Hierarchy	73
19	Developmental Analysis for Stoichiometry Hierearchy	78
20	Developmental Analysis for Molarity Hierarchy	81
21	Developmental Analysis for Food Web Hierarchy	84

LIST OF FIGURES

FIGURES		PAGE
1	Specimen data for transfer	15
2	Experimental design	22
3	The hypothesized stoichiometry hierarchy	31
4	Sub-hierarchies one to five for the stoichiometry hierarchy	35
5	The validated stoichiometry hierarchy	41
6	The hypothesized molarity hierarchy	44
7	Molarity sub-hierarchies	48
8	The validated molarity hierarchy	52
9	The hypothesized conservation of mechanical energy	53
10	Sub-hierarchies A and B for conservation of mechanical energy hierarchy	57
11	Inclusion validated conservation of mechanical energy hierarchy	59
12	Hypothesized and validated food web hierarchy	62
14	Sub-hierarchies for Mendel's laws Hierarchy	72
15	The inclusion validated Mendel's laws Hierarchy	74

CHAPTER ONE

INTRODUCTION

How to identify optimal instructional sequences is a long-standing and continuing concern of educators. As Posner and Strike (1976) indicate, a variety of approaches has been tried. Some of these may be categorized as "learning-related." The work described in the present report falls into this group. In particular, Gagne's learning hierarchy model (Gagne, 1970) was applied to six varied science concepts. Further, the relationship of learner developmental level, in the Piagetian sense, to learning of the same concepts was investigated. The particular concepts studied were "stoichiometry" and "molarity" from chemistry, "mendelian genetics" and "food webs" from biology, "conservation of mechanical energy" from physics, and "rock classification" from geology. The last named did not develop well enough to warrant major reference in this report. In each case the following research questions were addressed:

1. Can the development of the concept be represented in an empirically valid learning hierarchy?
2. What is the relationship between learner intellectual development and acquisition of the science concepts examined?

In the sections which follow, Gagne's learning hierarchy model and Piaget's developmental theory will be briefly described. It is assumed, however, that potential readers are familiar with the basic model in each case.

Theoretical Background

The fundamental premise underlying Gagne's learning hierarchy model is that, for some kinds of learning, it is possible to analyze a target outcome such that a series of subordinate capabilities may be identified, each of which is both necessary and useful to the learning of the capability immediately

following. The product of such an analysis is referred to as a learning hierarchy. Originally the hierarchical model was assumed to apply to all kinds of learning. Later it was suggested to be applicable to learning in only one of the five domains which Gagne suggests represent all learning. This domain is referred to by Gagne (1972) as the domain intellectual skills.

It is important, to clarify the meaning of the term intellectual skill in the sense implied by Gagne. As an illustrative example, suppose that an individual is learning about Boyle's law, which states that for a fixed mass of gas at constant temperature, the volume of gas is inversely proportional to its pressure. Correct application of the law in a meaningful way to a problem type which is not novel to the individual represents the use of an intellectual skill. The way in which the individual attacks a novel problem represents the use of particular cognitive strategies. A statement of the law itself represents the use of verbalized knowledge. Manipulation of apparatus in determining or applying the law involves the use of motor skills. Finally, the feelings of the individual as he does these things represent outcomes in the attitude domain.

The difference between intellectual skills and cognitive strategies may not be immediately apparent to the reader. In fact Gagne and Briggs (1974) describe cognitive strategies as very special kinds of intellectual skills, which are important enough and distinctive enough to warrant differentiation from other intellectual skills. Gagne and Briggs suggest that it is the object of the skill which differentiates cognitive strategies from other intellectual skills. Intellectual skills are oriented towards the individual's environment, such as the use of graphs or calculation of the number of atoms in a given mass of an element. In contrast, the essential characteristic of cognitive strategies is that they have as their objects the learner's own thought processes.

Interestingly, given the joint focus of the present study, Gagne and Briggs suggest a link between the domain of cognitive strategies and Piaget's stages of intellectual development, in that they claim that the individual's repertoire of cognitive strategies sets limits to the kinds of problem solving he can successfully perform and hence to his intellectual development as defined by Piaget.

Restriction of the learning hierarchy model to one of five domains of learning seems a substantial delimitation. However, it is apparent that this domain, relating primarily at the school level to discriminations, concepts and the application of rules relating to these, is of fundamental importance. According to Gagne, the single most important factor involved in the learning of an intellectual skill is to ensure that the learner has prior mastery of, and recalls, all skills considered to be immediately subordinate to the new skill to be learned. Intellectual development is seen as the progressive attainment and combination of a wide variety of intellectual skills. In this sense, according to Gagne (1968), learning precedes development.

A further restriction of Gagne's model is to much smaller units of content than originally advocated. Gagne (1973) suggests that units representing the content to be covered in a single lesson are most appropriate. While this may be desirable for reasons of experimental control in studies designed to identify learning hierarchies, it is not likely to be of use to curriculum developers and instructional designers. In the present study, units representing one to two weeks of classroom instruction were used.

Proponents of the learning hierarchy model argue that it is fundamental to the learning of intellectual skills. Gagne suggests that "the single most important factor involved in the learning of an intellectual skill is to ensure that the learner has mastery of, and recalls, all skills considered to be

subordinate to the skill to be learned" (Gagne, 1968, p. 177). A number of reviewers are supportive of this and Gagne's general position (Briggs, 1968; Cotton, Gallagher and Marshall, 1977; Jones and Russell, 1979; Resnick and Wang, 1969; Walbessar and Eisenberg, 1972; White, 1973). However, the model has not been particularly attractive to the science education community. There are several possible reasons for this, each of which is debatable. Not surprisingly, given its origins, the model is often seen as excessively associationist. Novak (1977, p. 125), for example, accuses Gagne of "an intellectual sleight of hand" in that he explains simple learning in terms of stimulus-response connections but ignores this in moving on to more complex learning types typically encountered in school settings. However, the sleight may merely be in Novak's own mind as he is correct when he suggests that Gagne does not attempt to explain the learning of concepts, rules and combination of rules in S-R terms. Indeed, Gagne has been more concerned with identification of appropriate conditions for learning and for the processing of information than with theoretical explana-tions.

Strike and Posner (1976) also attack the learning hierarchy approach in terms of its associationist connections. Further, they argue that emphasis on behavioral objectives may often result in confusion between the goals and the behavioral effects of having achieved these goals. However, it should be noted that most applications of the learning hierarchy model have been concerned with much smaller units of content than the original K-6 SAPA curriculum project, and have not been the starting point for delineation of major goals. Okey and Gagne's (1970) hierarchy for solubility product, White's (1974a) hierarchy relating to kinematics, Linke's (1976) hierarchy concerning graphical skills, Beeson's (1977) hierarchy for electrical skills and Griffiths' (1979) hierarchy

leading to performance of a core skill for the mole concept, all relate to content expected to be covered in several class periods. Selection of these topics appears to have been made not on the basis of the identification of major goals of education, but rather on the basis of much more narrowly defined skills which follow from consideration of the major goals of science curriculum. In each case the focus is on demonstration of learning of intellectual skills, not generation or assessment of goals. Behavioral statement of these skills is a necessity not a confounding influence.

A further criticism, typified by Johnstone (1979) and Lovell (1979) is that the adult reasoning represented in a logically derived hierarchy is not necessarily consistent with the reasoning patterns of the younger learner. By implication, the adult hierarchy may not facilitate learning for the younger learner. For hierarchies such as those listed earlier it is difficult to see how the expert-generated hierarchy may be circumvented. In these hierarchies, the terminal skill and each successive lower skill is essentially broken down to progressively simpler components. A pathway including, but not necessarily restricted to, what is logically acceptable to the expert must surely be substantially followed for eventual success. This does not deny the merit of the view that we should pay attention to the characteristics of the individual learner, including his personal misconceptions relating to the particular task(s). In this way, transfer of learning through the hierarchy may well be enhanced. As Case (1978) claims, such a procedure may especially benefit those in need of remedial attention.

An alternative, and indeed more widely applied, model of human learning is provided by Piaget's stage theory of intellectual development. In this model, human intellectual development is represented in four discriminable

stages, each represented by a wide variety of behaviors which are interpreted by Piaget to imply the existence of characteristic mental structures. For example, an individual who is considered to be concrete-operational is not bound by intuition, can reason logically about concrete objects and events, exhibits conservation of parameters such as area and weight, and can establish a one-to-one correspondence between two observable sets of events. An individual is considered to be formal-operational if he can exhibit these kinds of behaviours and in addition exhibits behaviours such as recognizing the need to control variables, reflective reasoning, proportional and analogic reasoning, propositional reasoning and combinatorial reasoning.

Assuming the validity of the stage model, it should be possible to determine the prevailing developmental level of an individual by testing on an appropriate set of Piagetian or related tasks. It should also be possible to analyze curriculum content in terms of its developmental stage requirements. These considerations lead to the attractive prospect of optimizing curriculum and instruction by matching learning experiences with the learner's stage or, at a finer level, sub-stage. This prospect was enthusiastically embraced by science educators throughout the 1970's and to a lesser extent since. Renner and Stafford, for example, typified this when they wrote, "All of the effort to develop scientific literacy is lost ...if students do not encounter concepts that are on their intellectual level (Renner and Stafford, 1979, p. 300).

An extensive body of literature exists which relates to each of the Piagetian and Gagnean models, including many review articles. It is not intended to further review them here. It is noted, however, that no consensus of opinion exists about either. Some researchers (e.g., Jones and Russell (1979)) argue for the Gagnean approach. Some, such as Karplus (1977), argue for the Piagetian

approach. Some, such as Bergan (1980) argue against the learning hierarchy model in the form originally proposed by Gagne. Some, such as Brainerd (1973), argue against the notion of developmental stages. Some, such as Novak (1977), argue against both models. Some, such as Strauss (1972) argue that the models are too disparate to be meaningfully combined. Others, such as Beilin (1970), suggest that they can. Still others, most notably Case (1974; 1978) begin from components of each model and arrive at a position radically different from either. With the exception of Griffiths (1979) there do not appear to have been any direct attempts to compare the usefulness of the two models as predictors of learning of the same concept. The present study attempts to do this for a number of concepts.

In the next chapter, methods used to validate learning hierarchies and to determine learner developmental level are discussed with particular emphasis on the techniques used in the present study.

CHAPTER TWO

EXPERIMENTAL METHODS - RELATED LITERATURE

Hierarchy Validation - Early Methods

Ideally, when one skill in a hierarchy is considered superordinate to another, it should not be possible for an individual to exhibit the superordinate skill unless prior learning of the subordinate skill has already occurred. Further, learning of the subordinate skill should result in transfer of learning to its superordinate. Each of these criteria is important, although typically the first has been most often used in learning hierarchy validations. In principle, establishing the validity of a hierarchy in these terms appears straightforward. In practice, it has proven more difficult. Prior to the introduction of the "test of inclusion" (White and Clark, 1973) the methods used could best be described as quasi-statistical. The most used methods were Gagne's positive transfer index (Gagne and Paradise, 1961); Eisenberg and Walbessar's (1971) consistency, inverse consistency, adequacy, inverse adequacy, necessity and completeness ratios; the phi coefficient (Capie and Jones, 1971); and reproducibility (Resnick and Wang, 1969). White (1974) clearly demonstrated the limitation of these methods, noting

> All are shown to be unsatisfactory for one or more of the following reasons: sometimes the index can have values which indicate a hierarchical connection even when the skills are really independent; errors of measurement are ignored; there is no way of calculating the probability of error in generalizing from sample to population; and subjective judgements are necessary on how many questions of a set must be answered correctly for a "pass" in a skill. (p. 65).

Hierarchy Validation - Recent Inclusion Methods

The White and Clark (1973) test of inclusion overcame the above limitations. It focused on the relationship between each pair of skills hypothesized to be hierarchically related, and, by combination of the relationships thus established, on the validity of the overall hierarchy. The method is probabilistic and, in combination with White's (1974b) criteria for appropriate experimental control, was a major advance in the techniques available.

Briefly, the method usually involves testing the skills in the hierarchy by two test items each. Ideally, for any pair of skills there should be no subjects with both items correct for the superordinate skill and both items wrong for the subordinate skill. Determination of whether the actual number of such subjects exceeds the number to be expected by chance, as calculated from the skill-by-skill response matrix, represents the basis of the White and Clark test. In order to determine the critical number of exceptions from the response matrix, assumptions must be made about the mastery status of those subjects who are neither clearly masters nor clearly non-masters of a particular skill. As Owston (1979) has shown, this procedure can, under some circumstances, lead to faulty parameter estimates.

A method which subsumes and extends the White and Clark test has been developed by Dayton and Macready (1976a). This method focuses on a number of skills simultaneously. According to Owston, this method suffers from the same limitations as the White and Clark test. However, this criticism may not be justified in that it is possible to eliminate intermediate cases from the analysis, thereby making the parameter estimates more robust. Further, the simultaneous consideration of a number of skills may also have the same effect.

Because this was the main statistical technique applied in the present study it will now be described in some detail.

Briefly stated, the probability of a subject producing a specific response pattern u, if the hierarchy is valid is given by

(1) $$P(u) = \sum_{j=1}^{q} P(u|v_j)\theta_j,$$

where

(2) $$P(u|v_j) = \prod_{j=1}^{k} \alpha_i^{a_{ij}} (1-\alpha_i)^{b_{ij}} \beta_i^{c_{ij}} (1-\beta_i)^{d_{ij}},$$

v_j represents the set of j consistent ("true") response patterns, and k is the number of skills in the postulated hierarchy. α_i and β_i represent the probability of a subject falsely exhibiting a skill and a subject falsely failing a skill, respectively, and are labelled guessing and forgetting parameters by Dayton and Macready.

Essentially what is implied by the use of Eqs. (1) and (2) is that the values of the misclassification parameters α_i and β_i are respectively raised to a power representing the number of guessing and forgetting corrections necessary to fit all true pattern vectors to an observed data vector. Similarly, $(1-\alpha_i)$ and $(1-\beta_i)$ are raised to the number of "correct" responses in each case. The product of these overall possible response patterns is Eq. (2). Multiplying this by the probability that the jth true pattern vectors, yields Eq. (1), which represents Dayton and Macready's general probabilistic model.

Dayton and Macready discuss the use of their model with respect to three cases. In Case A, α_i and β_i are unrestricted. That is, there are separate guessing and forgetting parameters for each task, although the same parameters characterize a given task across different "true" pattern vectors. Unfortunately, Case A has been solved to date only for concept attainment models, defined by

the authors as models in which each subject responds completely correctly or completely incorrectly to a given set of tasks. While of potential use to those interested in criterion-referenced testing, at present Case A is, therefore, not of general use in validating learning hierarchies. In Case B, all α's are the same for different tasks and all β's are the same for different tasks. In Case C, only one misclassification parameter is included (i.e., $\alpha_i = \beta_i = \alpha$). In all three cases maximum likelihood estimates of the various α's, β's and θ's are obtained through a series of iterations. These values are used to compute the number of expected responses for each possible response pattern. The goodness of fit between data and the hierarchy model is then calculated by both a Pearson chi-square test and a likelihood ratio expressed in the form of a chi-square. The latter appears to be more useful as it is less severely distorted by small frequencies.

In addition to the White and Clark test and the Dayton and Macready test, a test called the ordering-theoretic test (Bart and Krus, 1973; Airasian and Bart, 1975) has gained some prominence in recent years. Although likely to be less precise than the White and Clark test, the method produces similar results (Griffiths, 1979). As a sifting agent prior to use of the Dayton and Macready test it may serve a useful purpose. It has been used in this fashion in the present study. Basically, it is similar to the White and Clark test in that skills are tested in pairs. However, skills are tested directly from a pass-fail matrix rather than through an item response matrix. The proportion of subjects failing the lower skill and passing the upper skill must be less than a prespecified percentage for the connection to be considered valid. Normally, researchers using this method have allowed up to five per cent exceptions.

Hierarchy Validation - Transfer Methods

The limitations of Gagne's proportion positive transfer index have already been indicated. Yet the need to establish that transfer exists between related skills in a hierarchy remains. Several methods have been advocated in recent years. Each has some problems. The methods and their problems will now be considered.

White and Gagne (1974) suggest a direct test of transfer. For each pair of skills, two groups of subjects, all of whom possess neither skill, are formed. One group is taught the subordinate skill followed by the superordinate skill. The other group is taught the superordinate skill only. If the first group significantly outperforms the second, transfer is indicated. The procedure is repeated for each connection in the hypothesized hierarchy. White and Gagne's procedure has the advantage of being direct, but has important limitations. First, it may be unethical to deliberately expose subjects to a procedure which is anticipated to result in failure for half the sample, especially if the same subjects are to be used to test the relationships between a number of skills. Second, the procedure is unwieldy except for very small hierarchies.

Cotton, Gallagher and Marshall (1977) suggest what seems to be a theoretically more complex but empirically simpler version of White and Gagne's method, but the same general criticisms apply. In the same paper they also raise the question of whether substantial downward transfer as well as the upward transfer indicated by Gagne's model may occur.

Gagne' model is also criticized by Bergan (1980) who argues that it has several flaws which combine to reduce the amount of useful information which might be obtained for the relationship between intellectual skills. With

respect to transfer, Bergan claims that Gagne's model is inadequate in the following ways:

1. It does not allow consideration of transfer effects between skills which are not hierarchial psychometrically.
2. It ignores the existence of correlations between subordinate skills. Hence, what appears to be transfer from each of two or more skills to a common superordinate skill could be a function of correlational effects between the subordinate skills rather than transfer from each skill.
3. It ignores reciprocal causation.
4. It does not allow for transfer between non-adjacent levels.
5. It does not allow estimation of the relative strengths of transfer effects.

Bergan goes on to suggest that path analysis offers a means of testing for transfer, and can accommodate each of the above problems. Hence, according to Bergan, the determination of appropriate path coefficients allows estimation of each of the above effects as well as the direct transfer effects normally considered. This possibility would be clearly attractive to learning hierarchy researchers. Unfortunately, as will now be shown, the coefficients identified need not be related to transfer. Because of the currency and importance of Bergan's proposal it will be considered in detail.

The essential question addressed when the existence and magnitude of a learning transfer effect from one skill to another is assessed is whether learning the first skill first significantly enhances learning of the other. If it does, transfer is said to exist. In the simplest case of the relationship between two skills a substantial correlation between scores on tests of these skills, where one skill is also shown to be exhibited by less subjects, is consistent with the existence of transfer from the easier skill to the more

difficult one. However, it is not necessary nor sufficient as a means of establishing transfer. For example, consider the situations summarized in Figure 1.

Suppose that the data represented in Figure 1 have been collected for two skills, Skill 1 and Skill 2, respectively. The hypothesis under test is that Skill 1 is subordinate to Skill 2 and exhibits transfer to it. Figure 1(a) shows that of 200 subjects, 100 failed to exhibit either skill. The correlation coefficient between these two skills is high (.90). Given that Skill 1 is easier than Skill 2, the high value of r suggests good transfer from Skill 1 to Skill 2. However, suppose that the 100 subjects who failed both skills were now given a remedial treatment for the two skills, first Skill 1 and then Skill 2, and that the results represented in Figure 1(b) were obtained. All subjects gained Skill 1 but none gained Skill 2. In this circumstance, it does not seem possible to infer, contrary to the inference derived from the original correlation, that Skill 1 exhibits transfer to Skill 2. These data represent high correlation and low transfer.

Next, consider the data represented in Figure 1(c). In this case again 100 subjects fail both skills. However, the correlation coefficient in this case is close to zero. Again suppose that the 100 subjects failing both skills are given a remedial treatment. Further, suppose that in this case 50 gain neither skill and 50 gain both skills (Figure 1(d)). In a real sense, transfer from Skill 1 to Skill 2 is high. The penalty for failing to gain Skill 1, which is easier than Skill 2, is failure to gain Skill 2. Gain of Skill 1 for this group who failed to achieve it originally is followed by gain of Skill 2. The original correlation is low but transfer is high.

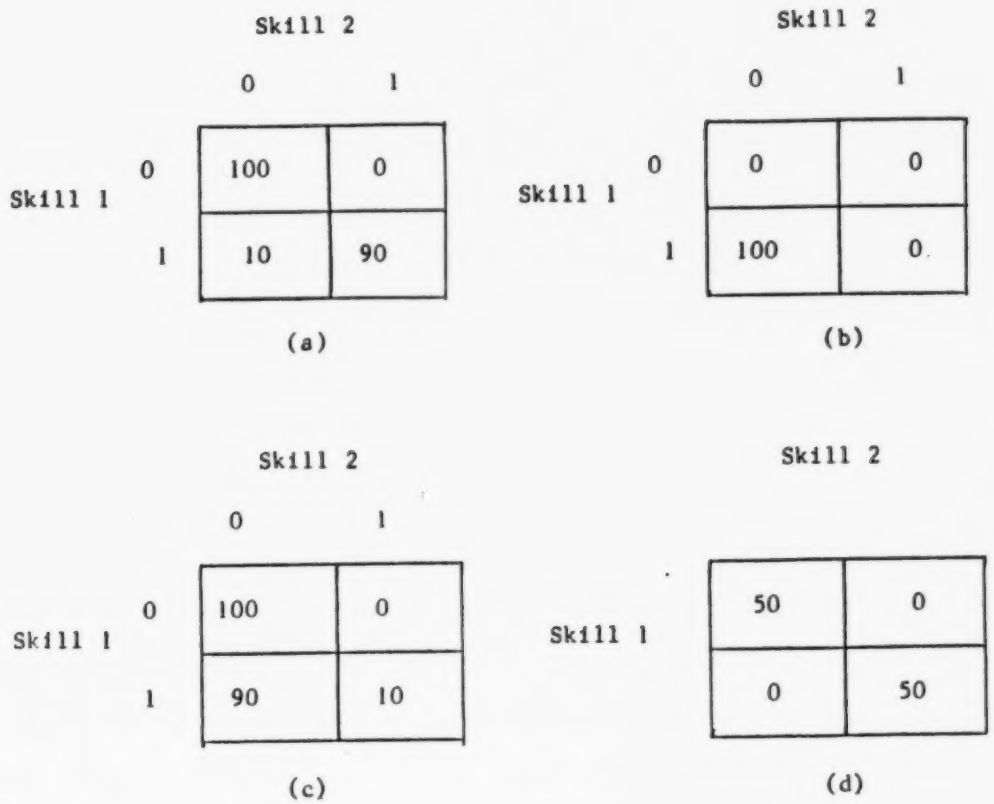


Figure 1 Specimen data for transfer

Returning to the question of the usefulness of path analysis as a means to assess transfer, it is suggested that the foregoing argument against the use of simple correlations for this purpose applies equally well to the use of path coefficients. Any skill considered to be superordinate to one or more other skills in a learning hierarchy will fall into one of the following categories: it may have only one subordinate skill; it may have two or more subordinate skills which are unrelated to one another; or it may have two or more subordinate skills at least some of which are related to one another. In each of the first two cases, the path coefficient from any particular subordinate skill to the superordinate skill equals the zero-order correlation between the two skills. For the data in Figure 1(a) a path coefficient close to unity would be obtained, suggesting excellent transfer. For the data in Figure 1(c) a low path coefficient would be obtained, suggesting poor transfer. However, as has been shown above, in the first case there need be no transfer and in the second case there might be. In each case, given the additional data in Figures 1(b) and 1(d), respectively, completely the wrong conclusion could be formed. Further, although the data used to illustrate this point are nominal, this need not be a problem (Boyle, 1970) and in any case similar results could be obtained starting with continuous data. When path coefficients are derived from appropriate combination of several correlation coefficients clearly the same conclusion would be reached, namely that path analysis is not a suitable technique to determine the degree of transfer between hierarchically related skills.

Each of the methods referred to thus far exhibits particular limitations, and one common limitation. In each case, each skill is tested once. Analysis is always correlational and indirect. In contrast, a method is now described which offers direct evidence of transfer between skills. The method

was originally proposed by Griffiths (1979, p. 196). Essentially, it focuses on the relationship between gain of subordinate and superordinate skills, respectively, between two parallel tests separated by a remedial treatment for skills evidenced as missing on the first test. Each test is post-instruction. The following steps are involved:

1. Each direct subordinate-superordinate relationship in the hypothesized hierarchy is identified.
2. Following instruction, a "pretest" is administered. (Note, the pretest is pre-remediation but post-instruction.) From this, those subjects who failed to learn the superordinate and subordinate skills in each relationship identified in step 1 are identified. These subjects represent the subsample for the particular connection involved. Generally, composition and size of these subsamples will vary between different connections in the hierarchy.
3. All subjects who failed to exhibit at least one subordinate-superordinate skill pair are directed to remedial work for the pair(s) of skill(s) initially failed.
4. Following remediation, each of these subjects is retested in a "posttest." The remaining steps are described for any pair of skills under test. The procedures involved are repeated for all hypothesized connections.
5. Subjects who exhibit the subordinate skill in the posttest are labelled "gain." Those who fail to exhibit the subordinate skill are labelled "no gain."
6. Similarly, for the superordinate skill subjects are labelled "pass" and "fail," respectively.

7. Transfer is assessed through a chi-square test of the strength of the relationship between gain/no gain and pass/fail.

This test clearly is still correlational. However, it is suggested that the correlation being tested represents a stronger test than that suggested by Bergan. Application of the test to the data in Figure 1 indicates no transfer in the first case and very strong transfer in the second. In each case, the conclusion is sensible in terms of the appearance of the data, in contrast to the conclusions reached from application of path analysis to the same data. The method is also preferable to that of White and Gagne in that there is no ethical objection because all subjects in need of remediation get it. In the present study the Griffiths method was therefore used as the test of transfer between skills in each of the hierarchies under consideration. Despite this it is recognized that the method has some limitations. In particular, if the number of subjects who fail either of the skills in a particular connection in the pretest is very small the test is not possible. Clearly, this is quite likely for the lower skills in a hierarchy. Yet, if the hierarchy is to be useful for instruction, the lower skills must be learnable by most of the population. Correspondingly, if the lower skills are readily learnable by most of the population, being able to show transfer of learning between those skills and from those skills to the intermediate skills of the hierarchy is not particularly useful. Conversely, it is important to identify skills which exhibit transfer of learning to the more difficult skills in the hierarchy. In effect, this means the middle and uppermost skills. For these skills, the Griffiths test appears to be the most suitable available, and was used throughout the present study.

Developmental Level - Group Testing

The educational and psychological literature relating to intellectual development, in the sense implied by Piaget, is extensive. In the present report, no attempt will be made to review this literature. Rather, one aspect of particular importance to this study, namely group-testing of developmental level will be referred to briefly. For a more extensive review the reader is referred to Nagy and Griffiths (1982).

Researchers concerned with the development of group tests of developmental level presumably make three assumptions. First, developmental stages as described by the Piagetian school, actually exist. Second, it is possible, by following interview procedures such as those of Piaget, to reliably determine the prevailing developmental stage of an individual. Third, it is possible to approximate this result for each of a group of individuals when they are tested as a group. Whether any or all of these assumptions are true has been a matter of considerable debate in the educational and psychological literature. However, the idea of being able to relate the intellectual development of learners to the intellectual demands of the curriculum content to be learned is appealing. To establish whether such a relationship exists it is necessary to be able to test large numbers of subjects. It is clearly much more economical in terms of the researcher's time to test these subjects in groups rather than individually. The same would be true for teachers wishing to apply the results of this research in their own teaching, at least if they wish to make the same kind of match.

The development of group tests of intellectual development is an area of research which is fraught with methodological difficulties. For example, it is difficult to assess the reliability of a test when, on the one hand, the use

of internal consistency measures such as KR-20 suffers from lack of agreement about how consistent responses to items testing different schema should be. On the other hand, test-retest methods may be influenced by memory effects in the short-term and by developmental growth in the long-term. Alternatively, testing itself may act as a teaching device, and hence inflate retest scores.

Studies by Lawson (1978), Renner (1977) and Shayer, Adey and Whylam (1981) are among the best conceived. In each of these studies validation of the group-administered instrument is made by correlation with the results of administering individual interviews, a procedure which is not always followed by all researchers but which seems absolutely necessary. Promising as these tests may be, the position was taken that none was adequate, and that a test should be specifically developed for the present study. The procedures are described in the next chapter.

CHAPTER THREE

PROCEDURES

In all, five hierarchies were completely investigated, and one (on rock classification) was investigated only to the stage of collection of pilot data. Data and analysis for each hierarchy are reported in chapter four. In effect, six independent studies were conducted, each with similar design and procedures. Design and procedures are described in the present chapter.

Design

The primary intent of the overall study was to attempt to identify valid learning hierarchies for a variety of science concepts chosen to represent different science disciplines. Hence, it was hoped, the application of the learning hierarchy model would be generalized substantially beyond its present primarily mathematics base. In most studies hierarchies have been validated either according to an inclusion definition or a transfer definition of hierarchial dependency, but not both. In the present study both perspectives were considered. A second purpose of the study was to determine the relationship, if any, between learner developmental level and learning of the intellectual skills comprising the hierarchy. The design, which was replicated for each hierarchy investigated, incorporates each of these facets of the study. It is represented in Figure 2. Actual sample sizes are given in Chapter Four in conjunction with the data for each hierarchy. However, about 200 subjects were involved in each sub-study. All, except for part of the genetics hierarchy sample, were from senior high school classes. In each case intact classes were involved. In each case a pretest and a posttest

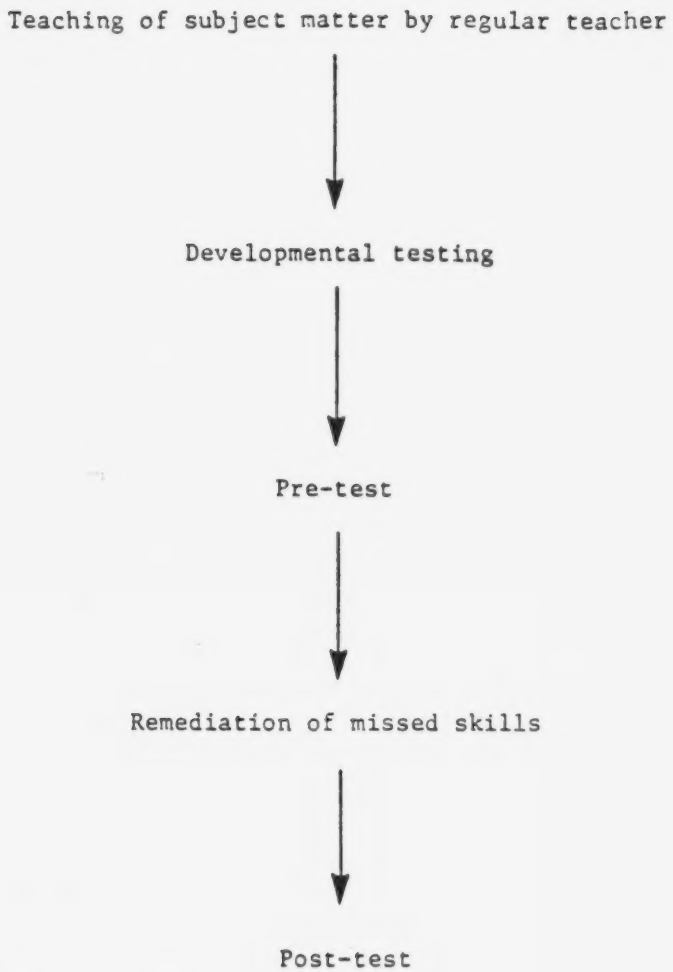


Figure 2. Experimental design

was administered to each individual subject in a written format. Almost all items were free-response. The pretest was administered after instruction on the topic had been completed by the classroom teacher. Pretest data were sifted according to the ordering-theoretic method initially, and then analyzed by application of the Dayton and Macready method. For analysis purposes the tests were marked carefully over an extended period of time. However, initially they were marked quickly and each subject was informed of his/her performance within a few days of taking the pretest. At the same time a remedial booklet was supplied to each subject, together with a prescription to study sections of the booklet representing skills which, according to the pretest, the subject had failed to gain during normal classroom instruction. Several days after this, the posttest was administered. In addition to the above tests a group-administered developmental test was administered to all subjects within the week prior to administration of the pretest.

Hierarchy Development

Common procedures were used to develop each hypothesized hierarchy. In each case the principal investigator, a research assistant, a graduate student and a science professor from an appropriate science discipline were involved. All had science and education degrees and had taught at the high school level. For each concept under consideration a target capability was chosen, such that it was representative of the essence of the concept and not beyond a level to which the concept is normally developed at the school level. The hierarchy was developed by asking of the target skill "what should the learner already be able to do in order to be able to learn this skill?", and then repeating this

question for each skill so derived, until a point was reached where skills obtained could reasonably be assumed to be already learned by the vast majority of learners. While this seems a one-step procedure, considerable dispute emerged for most of the hypothesized hierarchies. Hence, the generation of a hypothesized hierarchy acceptable to the team involved might best be described as an iterative procedure involving, typically, ten to fifteen iterations.

Hierarchy Test Development

For each hierarchy two parallel tests were developed. Each test was composed of a number of two-item subtests, one subtest for each skill in the hierarchy. The two items in any particular subtest were equivalent items. For each hierarchy, test items were devised by one member of the team, submitted to the team for scrutiny, and changed where necessary following discussion by the team. Items were then field-tested on a sample of about fifty subjects. This sample was as similar as possible in background to the sample to be used for the main study. Responses were examined both to identify technical problems in the items themselves and to determine whether items were equivalent. Changes were made accordingly and the tests revised. In each test, at the pilot and final testing stages items were arranged randomly. Because of their collective volume tests and remedial units are not included in this report. All are available from the author.

Remedial Unit Development

For each hierarchy a remedial unit was developed. The unit was necessary to the transfer design. The intent of the unit was to

promote learning of skills which subjects had failed to exhibit in the pretest. Hence, transfer could be tested by determining whether gain of superordinate skills was significantly related to gain of subordinate skills.

In each case the remedial unit was written by one of the team, and modified following consultation with the principal investigator. All units followed a common design. Behavioral statements of each skill were given, together with the hypothesized hierarchy. Learning activities for each skill in the hierarchy were presented in the order of the hierarchy. Each activity was clearly related to the skill involved, and incorporated test items similar to those used in the pretest and posttest. Answers to these test items were provided at the end of the booklet. Each unit was pilot-tested on a sample equivalent to the main sample. Because of their size these units are not included in this report but are available from the principal investigator.

Developmental Test Development

The literature relating to group testing of developmental level was referred to briefly in chapter two. In the present study, because of the large numbers of subjects needed to test the hierarchy, it was necessary to use a group-testing approach to the determination of individual developmental level. Several-field testing procedures were implemented first. Initially twelve widely-used tasks were administered by the principal investigator and a research assistant to a sample of 64 grade six students, 55 grade eight students and 51 grade ten students, all from the same St. John's school. For each of the six classes involved tasks were administered in three sittings to ensure

sufficient time for all tasks to be properly attempted. Each task was administered in a type-written booklet, and the investigator also explained the tasks verbally and made appropriate demonstrations.

The 12 tasks can be represented in four groups of three, representing conservation, proportional reasoning, combinatorial reasoning and controlling variables, respectively. The tasks were the displacement volume task (Karplus and Lavatelli, 1970), conservation of area (Goldschmid, 1967), conservation of weight (Piaget and Inhelder, 1962), ratio of distances (Wheeler, 1976), equilibrium in the balance (Inhelder and Piaget, 1958), a slightly modified version of the Mr. Short/Mr. Tall ratio task (Karplus and Karplus, 1970), combination in pairs (Piaget and Inhelder, 1975) a combinatorial task called the radio task (Hobbs, 1975), the pendulum task (Inhelder and Piaget, 1958), the shadows task (Inhelder and Piaget, 1958) and the spheres task (Wollman, 1977). Most of these tasks are well known and all have been referenced previously in the science education literature. In scoring these tasks subjects' explanations were used as the basis for scoring. For example, subjects' statements reflecting use of a combinatorial strategy were deemed necessary for a formal rating on any of the combinatorial tasks. Following this field-testing the combination in pairs task, the ratio of distances task and the shadows task were eliminated from further use. In the first year of the study the remaining tasks were administered to all subjects in the stoichiometry, genetics and conservation of mechanical energy sub-studies. However, because of problems in interpreting subjects' answers to the radio task and the balance task and because few high school subjects were unable to answer the conservation of area

and weight tasks these four tasks were eliminated from analysis. In the second year of the study only the remaining tasks were administered.

Although a group-format for administration of the developmental test had been used for the first year of the study, before collecting data in the second year it was decided to attempt to compare the results of administering the tasks in a group format and an individual format. Hence, a second pilot study was conducted. This study involved 100 subjects in grades eight and ten. The sample was divided into four groups representing intact classes. Two groups were tested by the principal investigator and two by a research assistant. For each researcher, one group was administered the developmental tasks in an individual interview format while the other group was administered the same tasks in a group format. After a gap of approximately one month, all subjects were retested in the other format. On each test administration, regardless of format, the volume conservation task and combinatorial task were answered correctly by substantially more subjects than were the controlling variables and ratio tasks. Overall difficulty indices were 0.49, 0.49, 0.27, and 0.18, respectively. The direction of these differences is consistent with the prevailing literature on these tasks, although the magnitude of the differences is a little surprising. In effect, two of the tasks may be considered to represent early formal thought while the other two may be considered to represent later formal thought. Despite this variability between items, there was considerable consistency in the results of administering the same items to the same subjects in the two different formats. As indicated in Table 1, for the sub-sample who were administered the group test

Table 1

COMPARISON OF GROUP AND INDIVIDUAL TESTING

Direction	N	% Change							r	sig.
		+3	+2	+1	0	-1	-2	-3		
G, I	50	2	6	28	56	8	0	0	.74	<.00
I, G	50	2	10	28	56	4	0	0	.79	<.00

Note: G, I = Group first
I, G = Individual first

first the Pearson product-moment correlation between individual and group test was 0.74. For the sub-sample who were administered the individual test first the correlation was 0.79. In each case, the correlation is substantial and significant at better than the .00 level. Further examination of Table 1 indicates that these correlations represent an underestimate of the strength of the relationship between the two test administrations because, regardless of which format was administered first, those subjects who changed in classification generally scored higher in the second administration. The contention that developmental tasks may be administered in a group format without undue loss of information compared to when the same tasks are administered individually is strongly supported by these data. It is also suggested that the test-retest information provided is an appropriate measure of reliability of the group test, given the inherent difficulty associated with direct test-retest and internal consistency estimates for Piagetian tests.

Statistical Tests

Each hierarchy was validated in terms of its inclusion characteristics by application of the Dayton and Macready test to the pretest data. Validation with respect to transfer was executed through the Griffiths test of transfer. The inclusion relationship between developmental level and individual's exhibition of intellectual skills with formal operational characteristics was investigated through application of the Dayton and Macready procedure.

Chapter Four

Hierarchy Validation

In this chapter data for each of the hierarchies under investigation will be described and discussed. In each case the degree of support for the hierarchy will be assessed. General conclusions and recommendations will then be formulated.

The Stoichiometry Hierarchy

When a chemist talks of stoichiometric relationships he is concerned with the amounts of reactants and products in a given chemical reaction. In effect, an understanding of stiochiometry allows the chemist to determine the amount(s) of product(s) produced when some particular amount of a specified reactant is consumed, or alternatively to determine the amount of reactant(s) needed to produce a particular amount of a desired product. Historically, it might be said that modern chemistry began with chemists' ability to do this. In a sense, it is possible to draw a parallel with the experience of today's chemistry students. When they have learned to use stiochiometric relationships, they have truly begun their understanding of chemistry. Stoichiometry is an important concept in introductory chemistry, and one which chemists use frequently.

The hypothesized hierarchy is represented in Figure 3. Each instrument was administered to intact high school chemistry classes in three schools over a period of about two weeks. After attrition the sample numbered 180 and contained 85 boys and 95 girls of mean age 15.4 years.

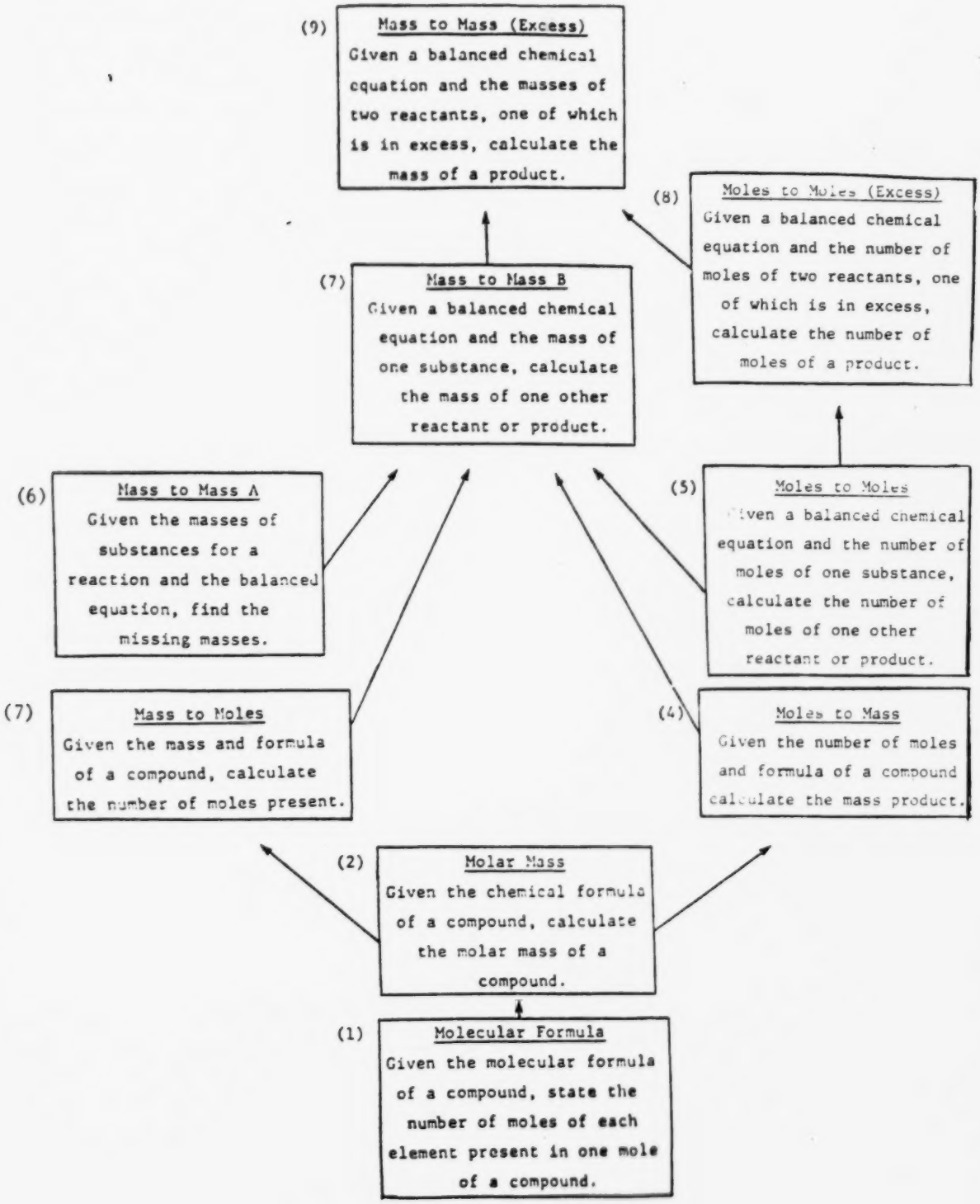


Figure 3. The hypothesized stoichiometry hierarchy.

Prior to analysis of the data, the reliability of each two-item subtest of the stoichiometry pretest and posttest was established through the extent of agreement in subject's responses to the two items. Level of agreement was high, as is indicated in Table 2.

Preliminary analysis of the data was made by applying the ordering-theoretic method. Percentages of exceptions to all direct and all indirect hypothesized hierarchical relationships are given in Table 3. Allowing five percent exceptions in each case, these data suggest the possibility that three direct connections out of 10 and four indirect connections out of 15 may be incorrect in the inclusion sense. In particular, Skill 8 may not be subordinate to Skill 9 and Skill 6 may not be subordinate to Skill 7, although it appears to be subordinate to Skill 9. Skill 2 appears to be equivalent to Skill 3 and Skill 4, rather than being subordinate to each of them. Finally, Skill 1 does not appear to be subordinate to Skills 2, 3 or 4. However, the remaining 18 direct or indirect connections are supported.

To further test both confirmations and anomalies a number of smaller sub-hierarchies were created, to which the Dayton and Macready test was applied. These sub-hierarchies and the results of the analyses are given in Figure 4 and Table 4. The combined hypotheses that Skills 2, 3 and 4 are subordinate to Skills 7 and 9, and that Skill 7 is subordinate to Skill 9 are supported by the Dayton and Macready parameters for sub-hierarchy 1. Similarly, the parameters for sub-hierarchy 2 support the hypotheses that Skill 5 is subordinate to each of Skills 7, 8 and 9. Further support for the relationship between Skill 5 and Skills 7 and 9 and between Skill 7 and Skill 9 is gained from consideration

Table 2

Hierarchy Sub-tests

Skill	Test Items		Sample Size		r (tet)	
	Pre	Post	Pre	Post	Pre	Post
1	1,1	1,1	163	165	.84	.98
2	7,6	6,7	163	164	.95	.98
3	4,2	2,4	163	164	.95	.98
4	2,5	5,2	163	165	.94	.98
5	5,9	9,5	158	164	.91	.99
6	8,8	8,8	161	164	.79	.91
7	6,3	4,6	163	164	.93	.93
8	9,7	7,9	152	161	.78	.93
9	3,4	3,3	163	164	.76	.81

Note 1: First test item in each case represents the item from the first half of the test.

Note 2: $p < .00$ for all pairs.

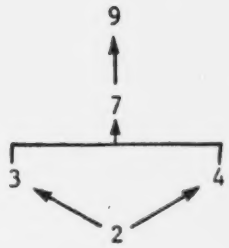
Table 3

Ordering-Theoretic Method: Percentage of Exceptions to Hierarchical Connections, Stoichiometry Pretest

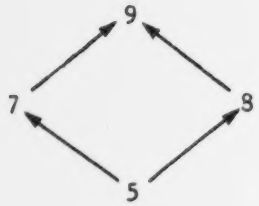
		"upper" skill								
		9	8	7	6	5	4	3	2	1
	9		1.3	10.6	12.6	25.2	31.1	34.2	36.6	16.0
	8	2.7		11.3		22.4	29.8			
	7	0.6	2.0		6.9	12.8	19.8	22.2	25.9	13.0
"lower"	6	0.6	0.7	5.6						
skill	5	1.3	0.0	1.3						
	4	0.0		0.0				2.5	2.5	0.6
	3	0.0		0.0			1.2		1.9	1.9
	2	0.0		0.6			0.6	0.6		
0.6										
	1	5.0		11.1			14.2	21.6	21.0	

Note: (1) The reader's understanding of this and similar tables may be aided by the following explanation. The percentage of exceptions for a particular hierarchical connection may be found by locating the "lower" skill for the connection and then reading directly across from it until the "upper" skill is located immediately above it.

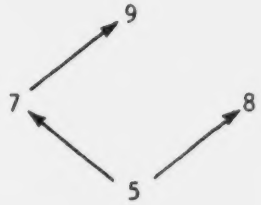
(2) A blank indicates that no hierarchical connection was hypothesized for the two skills concerned.



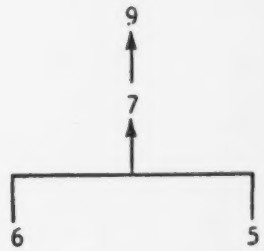
Sub-Hierarchy One



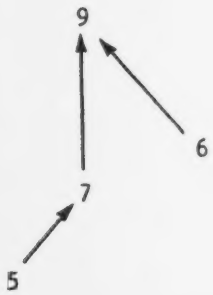
Sub-Hierarchy Two



Sub-Hierarchy Three



Sub-Hierarchy Four



Sub-Hierarchy Five

Figure 4. Sub-hierarchies one to five for the stoichiometry hierarchy.

Table 4

Dayton and Macready Analysis for Stoichiometry Sub-Hierarchies

Sub-Hierarchy	Guessing Parameter	Forgetting Parameter	Maximum Likelihood Estimate	Degrees of Freedom	Level of Significance
1	0.04	0.00	0.46	23	>0.99
2	0.00	0.05	3.91	8	>0.80
3	0.00	0.03	3.12	7	>0.80
4	0.09	0.03	8.85	8	>0.30
5	0.04	0.02	3.05	7	>0.80

of the results for sub-hierarchies 3, 4 and 5. In all of these cases the low values of the likelihood functions and the misclassification parameters support the original hypotheses.

The percentage of exceptions (5.6%) to the existence of a hierarchical connection between Skill 6 and Skill 7 casts some doubt upon the validity of the connection. Further analysis through application of the Dayton and Macready procedure to the data indicates that sub-hierarchy 4, a component of the overall hierarchy which includes this link, is consistent with the data at the .05 level of confidence. However, application of the same procedure to an alternative arrangement of the same skills represented in sub-hierarchy 4, namely sub-hierarchy 5, indicates that sub-hierarchy 5, which includes a link between Skill 6 and Skill 9 rather than between Skill 6 and Skill 7, is also consistent with the data. Further, the difference (X^2 = 5.80 , df = 1) between the likelihood functions for the fit of these hierarchies to the data is significant at the .05 level of confidence. An overall hierarchy with Skill 6 leading directly to Skill 9 appears to be preferable empirically to a hierarchy in which Skill 7 mediates the link between Skill 6 and Skill 9. However, the essence of Skill 6 is the understanding that mass is conserved in a chemical reaction. In terms of the chemistry involved, this is fundamental to the mass - mass relationships represented in Skill 7. Meaningful application of the relationships in Skill 7 clearly requires understanding of Skill 6, even if the empirical link between Skills 6 and 7 is more tenuous. Whether Skill 6 should be learned before Skill 7 is perhaps a matter of philosophical conviction of the value of meaningful learning, rather than of empirical evidence.

A second potential anomaly indicated by the ordering-theoretic analysis relates to the relationship between Skills 8 and 9. Sub-hierarchies 2 and 3 allow a test of whether Skill 8 is subordinate to Skill 9. As indicated in Table 4 each hierarchy is consistent with the data. Further, the difference between the likelihood estimates is insignificant at the 0.05 level. Hence, it does not seem to be essential to teach Skill 8 before Skill 9.

A third anomaly relates to Skills 2, 3 and 4. Meaningful learning of these skills seems to require prior learning of Skill 1. However, given that 21, 22 and 14 percent of the sample is able to exhibit skills 2, 3 and 4, respectively, without being able to exhibit Skill 1, it is clear that being able to state the number of moles of each element present in one mole of a compound is not necessary to the calculation of molar mass nor to the interconversion of mass and moles. Hence the hypothesized connections between Skill 1 and Skills 2, 3 and 4 are not supported by the data.

A final anomaly seems to occur for the relationships between Skills 2, 3 and 4. However, examination of the data indicate that few subjects failed to exhibit these skills. Hence, while they were clearly subordinate to all higher skills in the hypothesized hierarchy, the relationship between these three skills themselves could be obscured. To further investigate these connections, additional testing was carried out on a different sample, in this case grade nine students who had not been taught any of the content in the hierarchy. Two intact classes were randomly subdivided to yield four groups, A, B, C and D. Group A

(N = 12) received instruction on Skill 2 only and was then tested on Skills 2, 3 and 4. Group B (N = 14) received instruction on Skills 3 and 4 and was then tested on Skills 2, 3 and 4. All twelve subjects in Group A were successful on both items testing Skill 2. Five of these subjects were successful on all items for Skills 3 and 4 without instruction on those skills. In contrast, all 14 subjects in Group B failed to correctly answer any item testing any of the three skills. Further, when these subjects were then taught Skill 2 and retested on all three skills, 86 percent were successful on both Skill 2 items, 79 percent on both Skill 4 items and all were successful on the two Skill 2 items. Collectively, these results suggest not only that Skill 2 must be taught before Skills 3 and 4 but also that Skill 2 produces substantial transfer to Skills 3 and 4. Groups C (N = 15) and D (N = 16) were used to test the equivalence of Skills 3 and 4. Group C received instruction on Skills 2 and 3 only while Group D was instructed on Skills 2 and 4 only. Each was tested on Skills 3 and 4. All subjects were successful on the skills which they were specifically taught. In addition, only one subject failed to exhibit the untaught skill. Skills 3 and 4 appear to be equivalent.

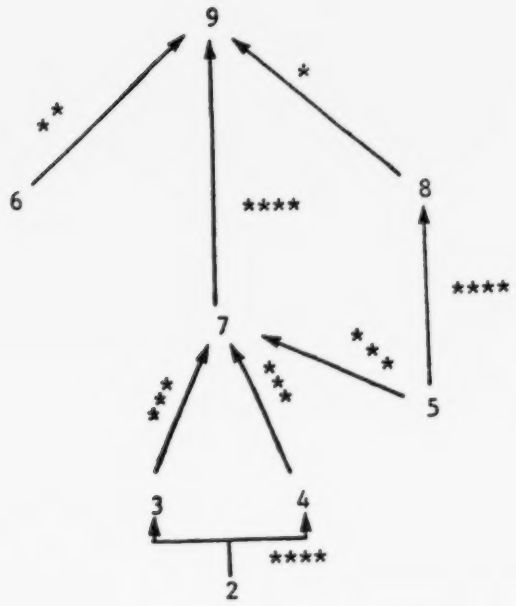
In addition to testing the inclusion validity of the hierarchy as described above, the Griffiths test of transfer was applied to the relationships involving Skills 5, 6, 7, 8 and 9. The test could not be applied to the connections involving the lower skills because too few subjects failed these initially. However, as indicated in the previous paragraph, there is good reason to believe in transfer from Skill 2 to Skills 3 and 4. For connections between Skills 6 and 7 and between

Skill 5 and Skills 7 and 8, respectively, Fisher's exact test was applied. The results are indicated in Table 5. They suggest good transfer from Skill 8 to Skill 9 and from Skill 7 to Skill 9. Skill 5 shows transfer to Skill 7. None of the other tests suggest significant transfer. These results are disappointing, and perhaps highlight the practical difficulty of establishing transfer. It is interesting to note that an understanding of the law of conservation of mass, a law which is fundamental to the existence of chemical equations, is not subordinate to the ability to perform calculations relating to these equations, either in an inclusion or a transfer sense. It is subordinate to the same skill in terms of inclusion when an excess of one reagent is employed, but again does not convey transfer. This suggests non-meaningful performance of such calculations. In summary of the above, a validated hierarchy, indicating transfer and inclusion characteristics is indicated in Figure 5.

The Molarity Hierarchy

The previous hierarchy dealt with stoichiometry. An integral part of that hierarchy was the concept of mole, for which a hierarchy has been reported elsewhere (Griffiths, 1979; Griffiths, Kass and Cornish, 1983). Mole and stoichiometry are, in turn, integral parts of the concept of molarity. Validation of a learning hierarchy for molarity represents the focus of the present section.

Molarity is a measure of the concentration of a solution. It is formally defined as the number of moles of a designated species which are present in exactly one litre of its solution. The species typically referred to are the original chemical in terms of its formula



Note: 1) * indicates that the connection is considered valid by the transfer definition of hierarchial dependency only.

2) ** indicates that the connection is considered valid by the psychometric definition of hierarchial dependency only.

3) *** indicates that due to the small number of subjects involved, the results obtained from the test of transfer for these connections cannot be meaningfully interpreted. Hence, although these connections are considered psychometrically valid, no decision can be made with respect to transfer validity.

4) **** indicates that the connection is considered valid by the transfer and psychometric definition.

Figure 5. The validated stoichiometry hierarchy

Table 5

Transfer Test, Stoichiometry Hierarchy

Connection	Subordinate Not Gained		Subordinate Gained			
	Subordinate Failed	Superordinate Passed	Superordinate Failed	Superordinate Passed		
8 to 9	8	1	2	13	13.2	<.00
7 to 9	11	1	1	8	13.6	<.00
6 to 9	3	0	5	6	2.86	>.05
6 to 7	2	0	4	5	1.85*	>.10
5 to 8	2	1	0	5	4.44*	<.05
5 to 7	2	0	3	2	1.12*	>.20

Note: For other connections N was too small to test. Nevertheless, as indicated in the test strong transfer was established from Skill 2 to both Skill 3 and Skill 4 by an alternate procedure.

2. * Fisher exact test applied.

before dissolving, or one or more of the ionic or molecular species present in solution after the original chemical has dissolved. For example, 1.00 litres of a solution containing 1.00 moles (111 grams) of calcium chloride ($CaCl_2$) may be alternatively described as a 1.00 M solution of calcium chloride, or as a solution which is 1.00 M with respect to calcium ions, or as a solution which is 2.00 M with respect to chloride ions, or as a solution which has a total ionic concentration which is 3.00 M. For different purposes, each of these kinds of interpretations may be appropriate.

The hypothesized hierarchy is represented in Figure 6. It is different from the other hierarchies considered in the present study in that it has three peaks instead of the usual one, although one of these is more prominent than the others. The two less prominent peaks are related to the third through common prerequisites. Collectively, the three branches thus produced in the overall hierarchy represent an overall understanding of the concept of molarity.

The general procedures for hypothesizing a hierarchy, production of tests and remedial unit, data collection and analysis, which were described in chapter three, were followed. Tests and remedial unit were field-tested on 41 grade ten students. Following minor changes resulting from this, data were collected on 144 subjects in intact high school chemistry classes.

Reliability of the two-item subtests was established through determination of a tetrachoric coefficient in each case. The coefficients obtained are represented in Table 6. In each case an acceptable level of agreement was obtained. Examination of the pretest and posttest

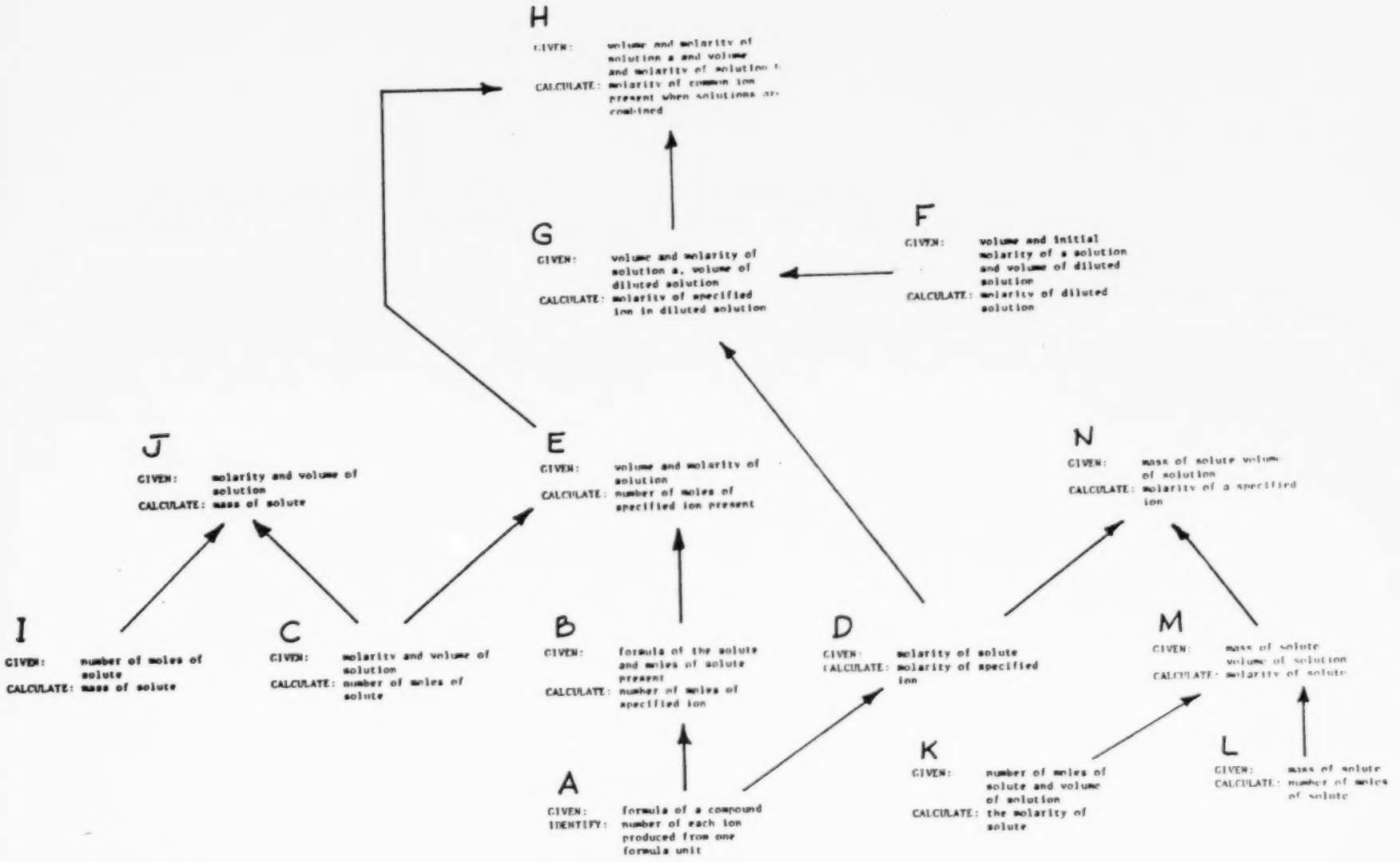


Figure 6. The hypothesized molarity hierarchy

Table 6

Tetrachoric Correlation Coefficient $r_{(tet)}$

Between Items Testing Skills of the Molarity Hierarchy

Skill	Corresponding test items		Sample Size		$r_{(tet)}$		p	
	Pre-	Post-	Pre-	Post-	Pre-	Post-	Pre-	Post-
A	9,26	(7,19)	127	(134)	.87	(.90)	<.001	(<.00)
B	11,27	(10,28)	124	(131)	.89	(.91)	<.001	(<.00)
C	2,28	(6,23)	124	(134)	.69	(.90)	<.001	(<.00)
D	3,23	(2,21)	130	(133)	.90	(.89)	<.001	(<.00)
E	5,16	(4,20)	134	(133)	.89	(.76)	<.001	(<.00)
F	6,21	(3,27)	131	(132)	.85	(.92)	<.001	(<.00)
G	10,24	(8,22)	129	(133)	.93	(.94)	<.001	(<.00)
H	14,20	(9,17)	131	(135)	.90	(.95)	<.001	(<.00)
I	13,17	(14,16)	134	(135)	.87	(*)	<.001	(-)
J	12,19	(5,18)	132	(134)	.90	(.91)	<.001	(<.00)
K	4,25	(1,24)	129	(134)	.50	(*)	<.001	(-)
L	1,18	(12,15)	132	(135)	.91	(.90)	<.001	(<.00)
M	7,15	13,26	134	(132)	.90	(.90)	<.001	(<.00)
N	8,22	(10,24)	130	(134)	.79	(.82)	<.001	(<.00)

Note: $r_{(tet)}$ value (*) not determinable where zero frequency occurs in one cell or more of the data matrix.

Student absences account for the difference in sample size.

data indicated a gradation of difficulty from lower to upper skills in general and between all skills hypothesized to be directly related, with the sole exception of the relationship between Skills H and G. Preliminary analysis was made by applying the ordering-theoretic test to the pretest and posttest data. The results of this are represented in Table 7. In general the results were supportive of the hypothesized hierarchy. The only exceptions to this were the relationships between Skill H and Skills G and F, respectively. In the case of these skills, the number of exceptions casts doubt on the hierarchical relationship in each case.

The existence of the hypothesized connections was tested more strongly by application of the Dayton and Macready procedure to the pretest data. Because of the size of the overall hierarchy this analysis was performed on a number of sub-hierarchies. These are represented in Figure 7. The results of the analysis are indicated in Table 8. Five of the eight sub-hierarchies tested in this way exhibited a good fit to the data ($p < .05$). Each of the exceptions again involved the hypotheses that Skills G and F were subordinate to Skill H. In effect this means that one direct hypothesized connection (Skill G to Skill H) and one indirect relationship (Skill F to Skill H) were not supported. However, the twelve remaining direct connections were all supported. In sum, 20 out of 22 hypothesized relationships were supported.

In addition to the inclusion relationships just discussed, transfer relationships between directly related skills were investigated in the same way as for the skills in the stoichiometry hierarchy. For seven of these connections the results are represented in Table 9. In

Table 7

Ordering-Theoretic Method: Percentage of Exceptions to Hierarchical Connections, Molarity Pretest

		"UPPER" SKILL													
		A	B	C	D	E	F	G	H	I	J	K	L	M	N
	A	-	3.2		1.1	0.0		1.0	0.0						0.0
	B	21.5	-			1.1		0.0	0.0						
	C			-		0.0		0.0	0.0		0.0				
	D	24			-			0.0	0.0						0.0
	E	43	17	56		-		2.0	0.1						
	F						-	2.0	6.7						
LOWER	G	59	31	73	26	14	28	-	6.8						
SKILL	H	62	32	74	23	11	31	3.4	-						
	I									-					
	J			4.8							-				
	K											-		0.0	0.0
	L												-	0.0	0.0
	M											5.8	7.6	-	0.0
	N	51			16							67	73	61	-

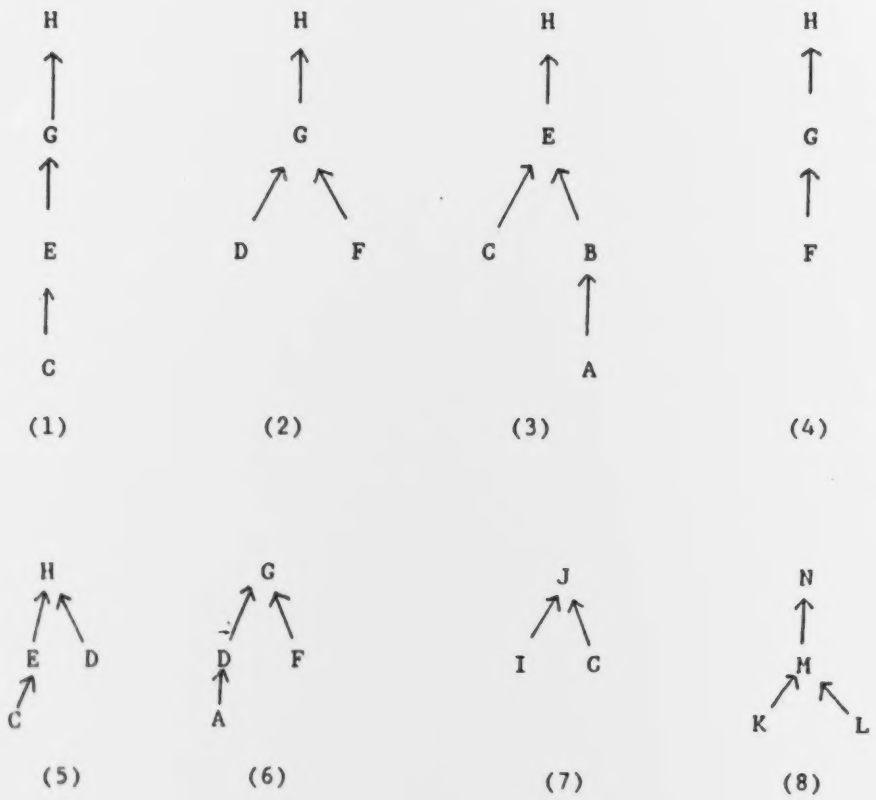


Figure 7. Molarity sub-hierarchies

Table 8

Dayton and Macready Analysis for Molarity Sub-Hierarchies

Hierarchy	Guessing Parameter	Forgetting Parameter	Maximum Likelihood Estimate	Degrees of Freedom	Significance
1	0.00	0.11	15.21	7	<.05
2	0.04	0.08	37.47	8	<.001
3	0.05	0.00	8.88	22	>0.99
4	0.04	0.09	6.89	2	<.05
5	0.00	0.01	2.56	7	>0.80
6	0.00	0.04	4.14	7	>0.30
7	Perfect fit, no exception				1.00
8					1.00

Table 9

Test of Transfer Molarity Hierarchy

Connection	Subordinate Not Gained		Subordinate Gained		X^2	p
	Superordinate Failed	Superordinate Passed	Superordinate Failed	Superordinate Passed		
G to H	26	2	19	18	11.01	<.00
E to H	18	0	15	15	10.87	<.00
F to G	13	0	6	19	16.83	<.00
D to G	9	0	9	19	9.98	<.01
D to N	8	0	6	17	10.28	<.01
B to E	3	0	7	15	2.67	NS
A to B	2	0	0	8		*<.05

NOTE: For all other connections, N was too small to test.

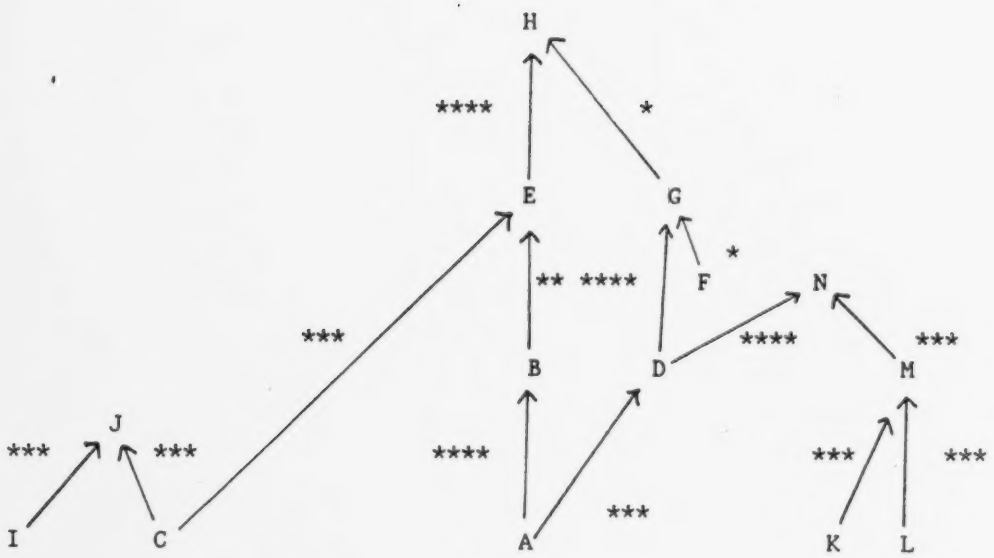
*Fisher test.

six of these cases (A - B, D - N, D - G, F - G, E - H, G - H) significant transfer was found. For the seventh connection (B - E), although the data were consistent with transfer, the effect was not statistically significant. For the remaining seven connections (A - D, C - E, C - J, I - J, K - M, L - M, M - N) too few students failed the lower skill in each pair to allow the test to be used.

The collective results of application of the inclusion and transfer tests are represented in Figure 8. The hierarchy represented in this figure is considered to be the validated molarity hierarchy.

The Conservation of Mechanical Energy Hierarchy

All objects have the capacity to do work on other objects. The energy which enables this to happen is referred to as the mechanical energy of the first object. This energy contains a component which is dependent on the position of the object, which is referred to as its potential energy, and a component which depends on the velocity of the object, which is referred to as its kinetic energy. The sum of these two components is constant. This relationship is referred to as conservation of mechanical energy. A hierarchy relating to this was hypothesized and is represented in Figure 9. Similar procedures to those described for the stoichiometry hierarchy were followed. After attrition, the sample was composed of 97 boys and 59 girls drawn from six grade ten classes in two schools. Content validity of each of the subtests of the pretest and posttest was established by expert opinion. The reliability of each subtest was established by determining a phi coefficient from the responses of the sample to the pair of items involved in each case. The results of this analysis are presented in



Note: 1) * indicates that the connection is considered valid by th transfer definition of hierarchial dependency only.

2) ** indicates that the connection is considered valid by the psychometric definition of hierarchial dependency only.

3) *** indicates that due to the small number of subjects involved, the results obtained from the test of transfer for these connections cannot be meaningfully interpreted. Hence, although these connections are considered psychometrically valid, no decision can be made with respect to transfer validity.

4) **** indicates that the connection is considered valid by the transfer and psychometric definition.

Figure 8. The validated molarity hierarchy

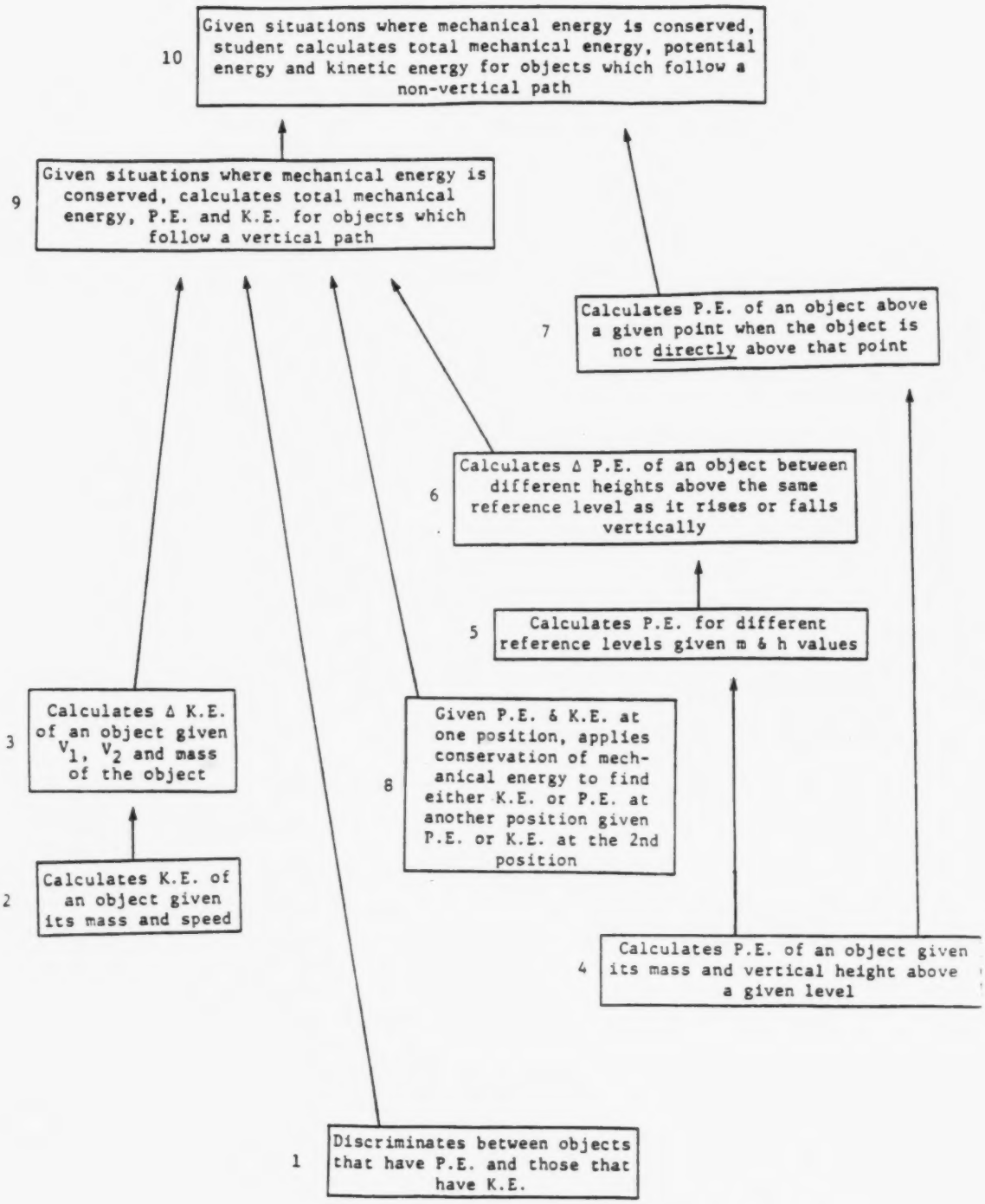


Figure 9. The hypothesized conservation of mechanical energy.

Table 10. All of the coefficients were significantly different from zero ($p < .001$) for the pretest data. For the posttest data, eight coefficients were significantly different from zero at the .001 level of significance, while the remaining two were significant at the .05 level of significance. Hence, each subtest was internally consistent.

Preliminary analysis by the ordering-theoretic method yielded the results indicated in Table 11. The general pattern of these results is consistent with the hypothesized hierarchy, except for the connections between Skills 2 and 3 and between Skills 5 and 6, respectively, and, marginally between Skills 7 and 10. Hence, Skills 3 and 6 were dropped from further analysis through the Dayton and Macready procedure, but the connection between Skills 7 and 10 was retained. For the application of the Dayton and Macready test the hierarchy, with Skills 3 and 6 removed, was subdivided into two smaller hierarchies. These are represented in Figure 10. The results, given in Table 12, indicate that the two sub-hierarchies are consistent with the data. Hence, the composite of these two hierarchies (Figure 11) was judged to represent the inclusion validated hierarchy.

As in the case of the stoichiometry hierarchy, it was not possible to test for transfer of learning from the lower skills in the hierarchy. The suggestion made in chapter two, that the lower skills in a hierarchy should be mastered by most subjects in a sample representative of the population for whom the hierarchy is intended, is reiterated here. Given the structure of the validated hierarchy, it is therefore reasonable to test for transfer of learning only from the intermediate skills, Skills 7 and 9, to the terminal skill, Skill 10.

Table 10

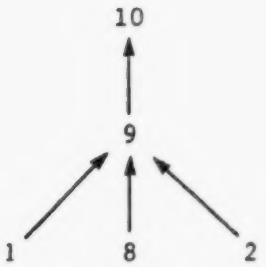
Phi Coefficients Between Items Testing the Same Skill on the Pretest, Mechanical Energy Hierarchy

Skill	Test Items	No. of Subjects	Phi Coefficient	Significance Level
10. Conservation of Mechanical Energy--nonvertical motion	9,18	149	.81	.00
9. Conservation of Mechanical Energy--vertical motion	6,14	147	.56	.00
8. Conservation of Mechanical Energy--basic application	3,16	149	.95	.00
7. Potential Energy--nonvertical reference point	5,11	149	.70	.00
6. Change in Potential Energy	7,19	148	.46	.00
5. Potential Energy--different reference levels	4,15	142	.55	.00
4. Potential Energy--basic application	2,17	152	.92	.00
3. Change in Kinetic Energy	1,20	148	.76	.00
2. Kinetic Energy--basic application	8,13	151	.94	.00
1. Discrimination between Potential Energy and Kinetic Energy	10,12	151	.38	.00

Table 11

Ordering-Theoretic Method: Percentage of Exceptions to Hierarchical Connections
Conservation of Mechanical Energy Pretest

		"upper" skill									
		10	9	8	7	6	5	4	3	2	1
	10		7.6	28.1	11.6	21.4	29.5	34.2	28.3	30.4	21.6
	9	3.5		16.7		13.3	17.3	21.1	16.8	18.5	14.4
	8	0.0	0.0								
"lower"	7	6.8						24.2			
skill	6	0.0	0.0				1.4	1.4			
	5	0.0	0.0			0.7		1.4			
	4	0.0	0.0		0.0	0.0	0.0				
	3	0.0	0.0							0.0	
	2	0.0	0.0						0.0		
	1	0.0	0.0								



Hierarchy A

Hierarchy B

Figure 10. Sub-hierarchies A and B for conservation of mechanical energy hierarchy

Table 12

Dayton and Macready Analysis for Kinetic-Potential Energy Sub-Hierarchies

Hierarchy	Guessing Parameter	Forgetting Parameter	Maximum Likelihood Estimate	Degree of Freedom	Significance
A	.00	.02	14.7	20	.70
B	.00	.04	21.4	22	.50

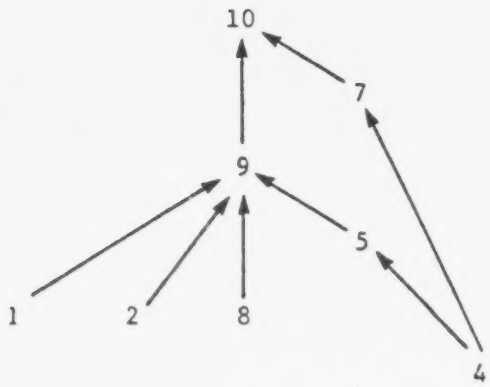


Figure 11. Inclusion validated conservation of mechanical energy hierarchy

Application of the Griffiths method yielded the results indicated in Table 13. Significant transfer is indicated in each case, offering further support for the hierarchy in Figure 11.

Before leaving consideration of the structure of the conservation of mechanical energy hierarchy, a further comment is necessary. As has been indicated already, the relationship between Skills 7 and 10 was marginally hierarchical. Examination of the responses of all subjects who were exceptions to the hierarchical link indicated that all had made the same mistake on the Skill 7 test items. In these items, in order to determine whether subjects understood that potential energy is related to the vertical height above a reference level, redundant non-vertical distances were included in the items as well as vertical height. The same kind of redundant information was not provided in the Skill 10 items. Normally, it is not provided for either kind of item. In this case, it seems likely that providing this information for items testing one skill but not the other may have produced spurious results. Further data were collected to test this. A sample of 123 subjects from two schools not used in the original study was used. Items representing Skill 10 were modified so that, like the Skill 7 items, the slope distances were given. No exceptions were found to the relationship between Skill 10 and Skill 9, and the connection between Skill 10 and Skill 7 was much stronger than for the original data, with only 1.7 percent exceptions. This finding not only indicates support for the hypothesized connection but also suggests the important implication that the normal practice of providing only necessary information in test items may obscure students' misconceptions.

Table 13

Test of Transfer for Conservation of Mechanical Energy Hierarchy

Connection	No Gain		Gain		X^2	df	Significance
	Fail	Pass	Fail	Pass			
7 to 10	3	0	4	18	5.18*	1	<.05
9 to 10**	6	2	0	12			<.00

* Yates correction applied.

**Fisher's exact test.

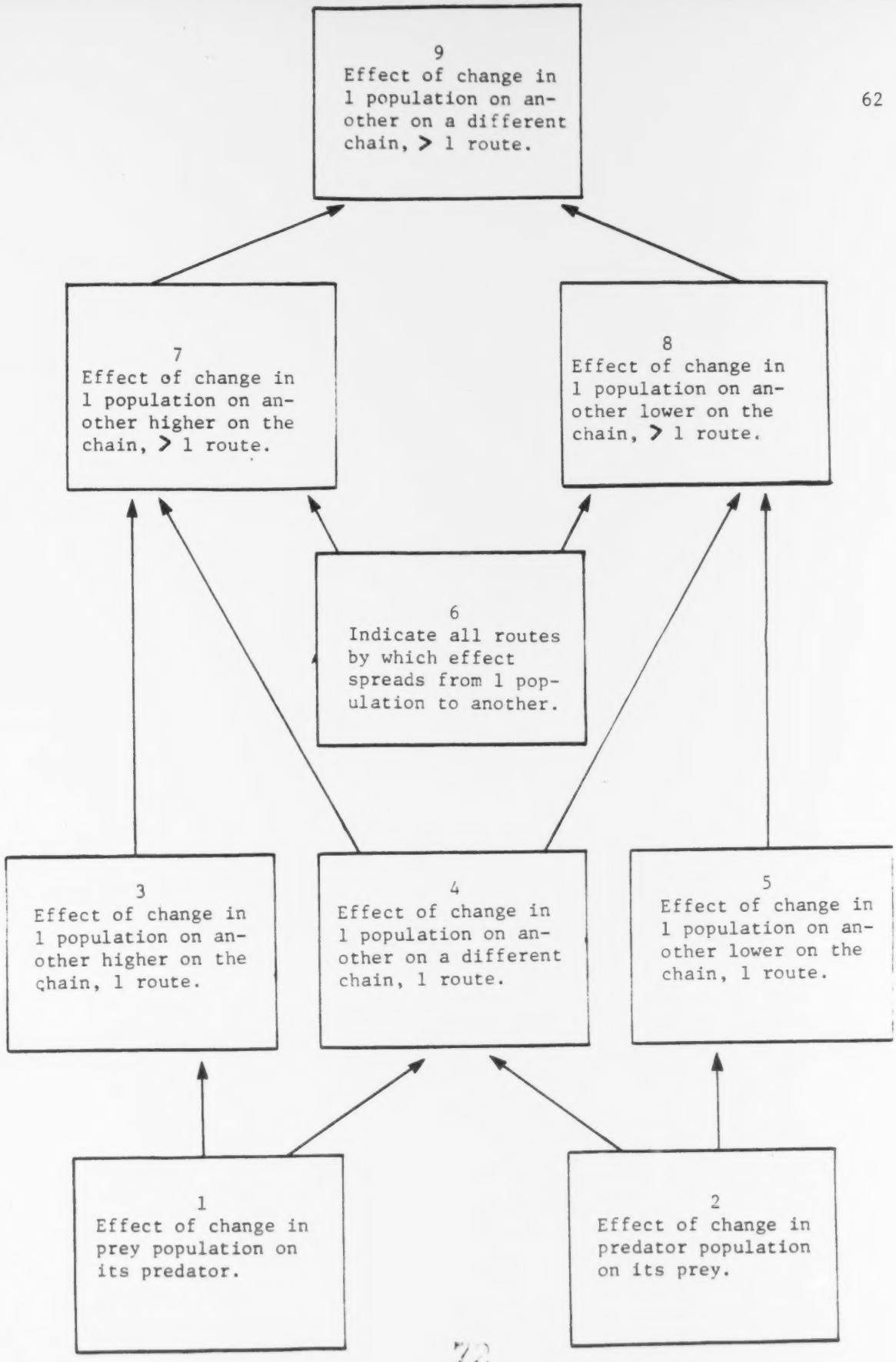


Figure 12. Hypothesized and validated food web hierarchy

The Food Web Hierarchy

The terminal skill of the hypothesized hierarchy was defined as "Given a food web diagram, determine the effect of a sudden size change in one population on a second population which is not on the same food chain, when the effect may be transmitted along more than one route." Subjects were not required to consider food webs involving more than four routes between populations in question. It was felt that increasing the complexity of food webs beyond this level would only increase the tediousness of the task, without demanding greater understanding of the basic skill. The hypothesized hierarchy is represented in Figure 12.

Procedures for formulation and testing of the hypothesized hierarchy were similar to those in each of the other studies. After a number of modifications, following discussion by the team of four researchers involved and subsequent field testing, the pretest, remediation and posttest were administered to 200 grade ten biology students in three St. John's schools. Eight coeducational classes and three teachers were involved. Pretest and posttest were in effect each composed of nine two-item subtests. Reliability of each subtest was established by determination of the tetrachoric correlation coefficient between the pair of items representing the skill involved. The results are given in Table 14.

In the case of the data for the food web hierarchy, some problems emerged which were not evident for the data for the other hierarchies. The tetrachoric correlation coefficient cannot be determined if one or more cells in the skill matrix has a frequency of zero. As a result,

Table 14

Tetrachoric Correlation Coefficients Pre- and Posttests

Food Web Hierarchy

	Skill	Corresponding Test Items	Sample Size (n)	r_{tet}	
Pretest	1	2,13	149	.85	<.05
	2	5,18	161	*	
	3	6,15	135	.62	<.01
	4	10,16	118	.79	<.00
	5	4,7	163	.72	<.00
	6	12,17	78	.87	<.00
	7	1,14	169	*	
	8	9,11	182	*	
	9	3,18	116	*	
Posttest	1	4,11	164	*	
	2	2,17	135	*	
	3	5,18	123	*	
	4	6,10	164	.75	<.00
	5	1,16	143	.89	<.01
	6	7,14	130	.98	<.00
	7	3,12	161	.95	<.00
	8	9,13	157	.95	<.00
	9	8,15	149	.75	<.00

*r_{tet} could not be determined for items for which zero frequencies occurred in at least one cell of the data matrix.

tetrachoric coefficients could not be determined from the pretest data for Skills 7, 8 and 9 because none of the sample correctly answered items testing these skills. However, tetrachoric coefficients were determined from the posttest data for parallel items testing these skills and were found to be highly significant. Skills 1 and 3 were mastered by most students on the posttest. As a result, zero frequencies occurred in at least one cell in each of the corresponding matrices. Although tetrachoric coefficients could not be determined for these skills in the posttest, parallel items were found to correlate well in the analysis of pretest data. Both items testing Skill 2 were answered correctly by 97.5% of the pretest sample and 98.5% of the posttest sample, respectively. As a result, tetrachoric coefficients could not be determined at all for items testing this skill. All other tetrachoric coefficients indicated a high degree of correlation between the two items testing a given skill.

In anticipation that some exceptions to the hypothesized hierarchy would be observed, it had been intended to apply the Dayton and Macready (1976) scaling procedure to the data. However, only one subject was found to be an exception to any of the hierarchical connections hypothesized. Hence, for the particular data obtained, the ordering-theoretic method is an appropriate test of the psychometric validity of the hypothesized hierarchy. Ideally, to prevent bias from exposure to the hypothesized hierarchy during use of the remedial booklet, only the pretest data should be used for such analysis, although this effect would probably be slight. In the present case, this was not entirely possible because no subjects passed the upper three skills on the

pretest, making tests of connections between these skills impossible. Hence, relationships between Skills 7, 8 and 9 were evaluated from the posttest data. All other connections were evaluated from the pretest data. These data are presented in Table 15. They show no exceptions to the hypothesized connections. The data are consistent with the hypothesized hierarchy.

The test of transfer was applied only to Skills 6, 7, 8 and 9 because too few subjects failed the lower skills in the pretest for the sample size to be considered large enough to apply the test of transfer. The results are presented in Table 16. Significant transfer ($p < .00$) was observed from Skill 6 to Skill 7, Skill 6 to Skill 8, Skill 8 to Skill 9, and Skill 7 to Skill 9. In effect, any connection for which it was feasible to test for transfer was supported. The hypothesized hierarchy is strongly supported by the data.

The Mendel's Laws Hierarchy

When living things reproduce some characteristics of the offspring may be related to characteristics of their forebears. In such cases the distribution of these characteristics among a group of offspring can be explained or predicted through application of Mendel's laws of inheritance. For example, suppose that a tall pea plant is "crossed" with a short pea plant. Suppose, further, that the offspring are all tall. If these offspring are then "crossed" three-quarters of their offspring will be tall and one-quarter will be short. This result would be predicted by application of Mendel's laws. Knowledge and application of these laws has long formed an important part of introductory biology courses, particularly at the high school and early

Table 15

Ordering-Theoretic Method: Percentage of Exceptions, Pre- and Posttest, Food Web Hierarchy

		Upper Skill								
		1	2	3	4	5	6	7	8	9
Lower Skill (Pretest)	1	-		0.0	0.0			0.0	0.0	0.0
	2		-		0.0	0.0		0.0	0.0	0.0
	3	2.1		-				0.0		0.0
	4	4.5	2.1		-			0.0	0.0	0.0
	5		13.3			-			0.0	0.0
	6						-	0.0	0.0	0.0
	7	97.8	98.5	93.8	93.3		20.6	-		0.0
	8	97.8	97.3		93.1	78.8	20.9		-	0.0
	9	97.8	98.9	94.9	92.3	75.0	19.0	0.0	0.0	-
Lower Skill (Posttest)	1	-		0.0	0.0			0.0	0.0	0.0
	2		-		0.0	0.0		0.0	0.0	0.0
	3	2.7		-				0.0		0.0
	4	3.7	1.8		-			0.0	0.0	0.0
	5		3.3			-			0.0	0.0
	6						-	0.9	0.9	0.0
	7	32.1	35.7	25.3	27.7		12.0	-		0.0
	8	34.9	35.4		28.8	22.6	11.1		-	0.0
	9	74.8	74.2	63.9	65.7	68.0	43.5	28.0	25.8	-

Table 16

Test of Transfer for Food Web Hierarchy

Connection	Subordinate Not Gained		Subordinate Gained		χ^2	
	Superordinate Failed (n)	Superordinate Passed (n)	Superordinate Failed (n)	Superordinate Passed (n)		
6 to 7	10	0	2	18	18.91	<.00
6 to 8	10	0	2	18	18.91	<.00
7 to 9	26	0	15	15	15.30	<.00
8 to 9	24	0	14	16	15.72	<.00

Note: all subsamples (n) are taken from those subjects failing both skills of a particular connection in the pretest.

college levels. In the present section the validation of a learning hierarchy leading to correct application of these laws is described.

The hypothesized hierarchy is represented in Figure 13. Similar procedures to those used for the other two hierarchies were followed. After attrition, the sample consisted of 221 boys and girls in approximately equal numbers. In this case, the sample was drawn from Grades 6, 9 and 11. The Grade 6 subjects were not tested on the upper three skills.

Preliminary analysis by the ordering-theoretic method yielded the results shown in Table 17. The general pattern of results was again fairly consistent with the hypothesized hierarchy. However, in this study much more variance between items testing the same skill was observed. The results should therefore be treated with caution, although it is noted that for purposes of analysis only subjects with both items right or both items wrong for any particular skill were retained in any analysis including that skill.

Skills 1, 2, 3 and 7 were found to be easy, even for the grade 6 students, and it was not possible to discriminate between these skills. However, all were clearly subordinate to the other skills in the hierarchy. Examination of the relationships between Skill 8 and Skills 11, 14 and 15, respectively, which were all hypothesized to be superordinate to it, suggested that none of these relationships were valid. Further consideration of Skill 8 itself and of the nature of subjects' responses to the test items confirmed that the skill was more difficult than originally believed and further that it was not a necessary component of the hierarchy. Hence it was not included in further analysis.

After consideration of the results of the ordering-theoretic

Table 17

Ordering-Theoretic Method: Percentage of Exceptions to Hierarchical Connections, Mendel's Laws Hierarchy

		"upper" skill														
		15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
	15	–	7.9	18.0	30.7	25.8	42.0	39.1	4.5	38.9	14.8	18.0	13.8	36.5	35.3	42.0
	14	1.1	–	6.4	20.4	13.8	26.9	27.2	3.2	26.7		6.4	9.8	25.6	26.7	29.0
	13	0.0	2.1	–	14.0	6.4		16.3	0.0	16.0		1.1	3.3	13.3	15.6	19.4
	12	0.0	0.0	1.1	–					2.6	0.5			2.0	2.8	3.7
	11	0.0	0.0	0.0		–			0.0	8.9				9.9	11.2	13.2
	10	0.0	0.0				–	0.0		2.6				1.5	1.8	2.7
"lower"	9	0.0	0.0	0.0			0.9	–		2.6				2.5	1.9	2.7
skill	8	3.4	12.9	18.3		17.1			–	38.7				33.3	37.4	41.7
	7	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	–				1.1	0.5	1.0
	6	1.1			30.7						–	5.6	12.1	35.6	41.6	44.3
	5	0.0	0.0	1.1							.9		5.1	15.0	19.2	18.4
	4	1.1	5.4	4.3							4.2	6.0		26.5	29.6	30.9
	3	0.0	0.0	0.0	0.5	0.0	0.5	0.0	0.0	0.6	0.0	0.0	0.0	–		1.5
	2	1.2	1.1	1.1	0.5	0.5	0.9	0.9	0.5	0.5	0.0	0.5	0.0		–	1.4
	1	0.0	1.1	1.1	0.5	0.5	0.5	0.5	0.5	0.5	0.0	0.5	0.5	0.5	0.5	

analysis, a number of sub-hierarchies, labelled F to O, were further tested through the Dayton and Macready procedure. These sub-hierarchies are represented in Figure 14 and the results of the analysis are given in Table 18. With the exception of Sub-Hierarchy H, all of the sub-hierarchies exhibited extremely good fit to the data. In the case of Sub-Hierarchy H, nine percent of the response patterns were invalid and the likelihood function was consequently large, indicating a poor fit to the data. The Dayton and Macready method offers comparison of alternative hierarchies containing the same skills. Comparison of Sub-Hierarchy H with Sub-Hierarchy I, suggests that Skill 4 (given the genotype of parents, list the genes which each parent could pass on to an offspring) is not subordinate to Skill 5 (given the genotype of parents, list all possible genotypes of offspring). No explanation is apparent for this, but the same result was observed on the posttest. Comparison of Sub-Hierarchies F and G also suggests a similar pattern, although the difference is not statistically significant. The only other comparison of special note is between Sub-Hierarchies N and O. Comparison of these two sub-hierarchies allows a test of whether Skill 10 is superordinate to Skill 9. No difference was observed. Comparison of the response patterns themselves suggests that these two skills are not hierarchically related. Consideration of the overall results described suggests that the hierarchy in Figure 15 is valid in the inclusion sense, although it is worth reiterating that the data for this concept is the most questionable in the study.

As in the case of the other hierarchies each direct connection in the hypothesized hierarchy was investigated for transfer. None was

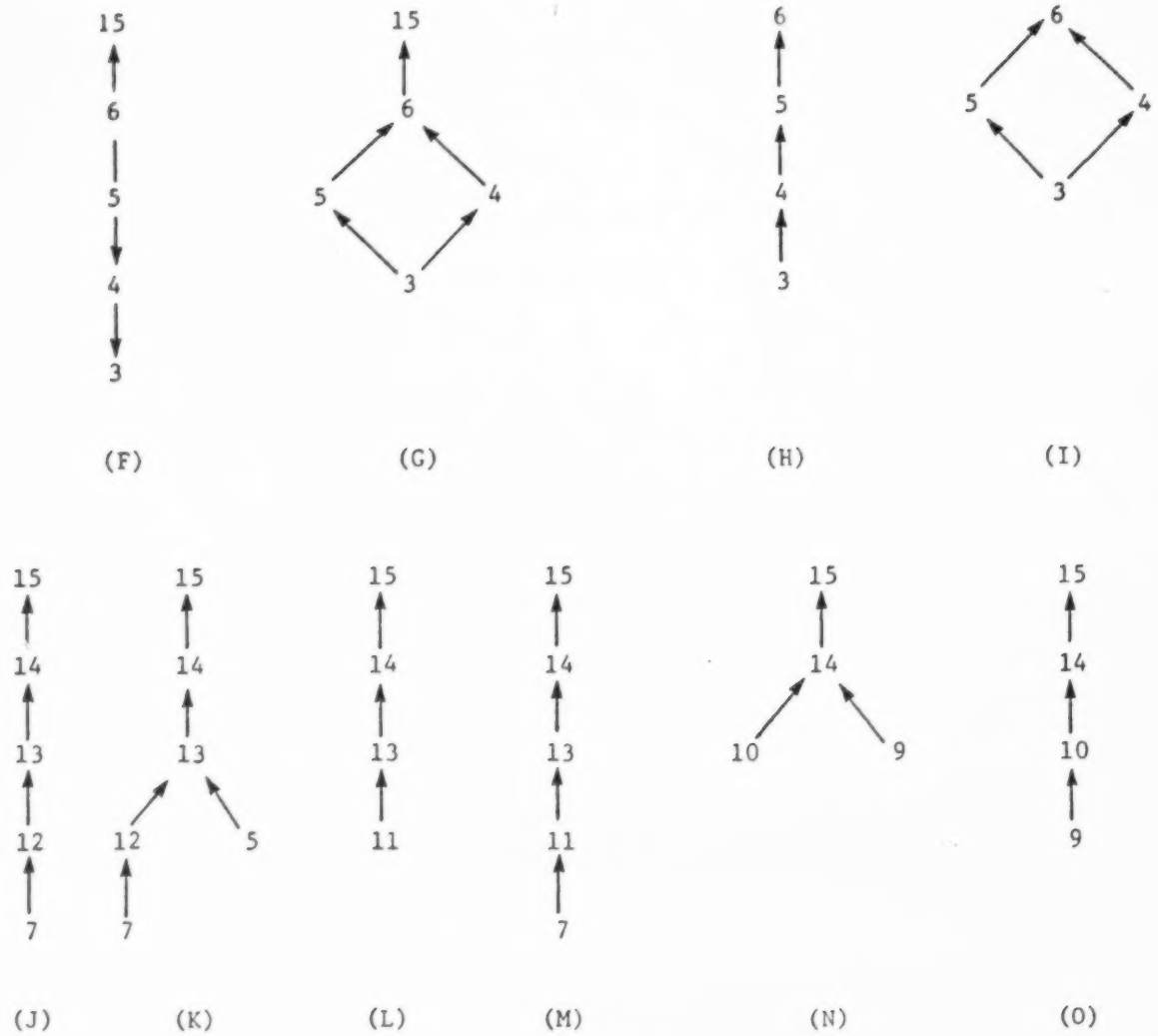


Figure 14. Sub-hierarchies for Mendel's laws Hierarchy

Table 18

Dayton and Macready Analysis for the Mendel's Laws Hierarchy

Hierarchy	Guessing Parameter	Forgetting Parameter	Likelihood Function	Degrees of Freedom	Significance
F	.25	.00	6.93	24	.99
G	.17	.00	5.24	23	.99
H*			*		.00
I	.00	.02	9.89	8	.29
J	.10	.00	12.04	24	.98
K	.00	.00	0.00	54	1.00
L	.00	.00	0.00	9	1.00
M	.00	.00	0.00	24	1.00
N	.05	.00	0.00	8	1.00
O	.05	.00	0.00	9	1.00

*Likelihood function extremely large.

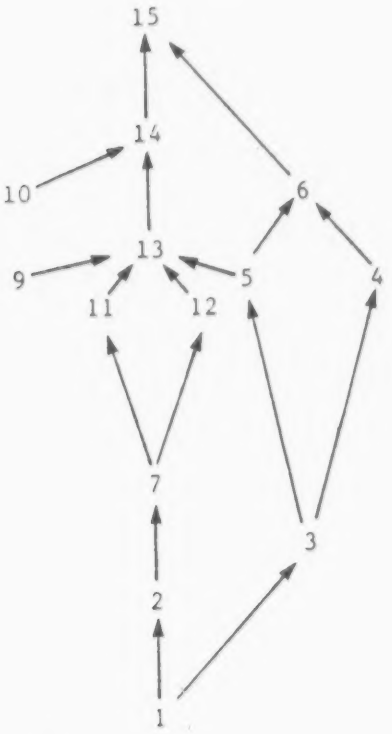


Figure 15. The inclusion validated Mendel's laws Hierarchy

observed. Few subjects gained at all from pretest to posttest. There are several possible explanations for this. Clearly doubt is cast upon the validity of the hierarchy. It may be that the hierarchy was deficient in the size of the steps between skills In other words, perhaps other skills intermediate between skills represented in the hierarchy should have been included. Another alternative is that the topic was not interesting enough for the subjects involved to wish to take advantage of the remediation. A third possibility is that the written format of the remedial unit was not suitable for the particular students involved. Unlike in the other studies, most of the subjects involved were from junior high or elementary school grades, and it is possible that the somewhat voluminous written remedial unit was unattractive to them. To further compound this problem, after the data were collected it transpired that the high school students in the sample were not considered academically well motivated by their teachers. As only these high school students were tested on the upper three skills in the hierarchy the low motivation of these students might be expected to inhibit any observed transfer. Hence, this hierarchy merits further investigation.

Summary

In the present chapter, data relating to the validation of a learning hierarchy for each of five concepts have been presented. In each case a hierarchy which appears to be reasonably valid in an inclusion sense has been identified. Only in one case, that of the hierarchy relating to Mendel's laws, was the hierarchy not also well substantiated in terms of its transfer characteristics. The relationship between acquisition of the more complex skills represented in these five hierarchies and learner developmental is described in the next chapter.

Chapter 5

Requirement for Formal Thought

In this chapter evidence is presented relating to the question of whether the learner needs to be capable of formal operational thought in order to be able to exhibit the upper skills in the hierarchies for which validation data was presented in chapter four. For analysis purposes subjects who performed three or four of the four developmental tasks correctly were considered formal. Those with zero, one or two tasks correct were considered non-formal.

The Stoichiometry Hierarchy

The terminal skill in the stoichiometry hierarchy requires solution of items such as "Consider the reaction between calcium fluoride (CaF_2) and hydrogen chloride (HCl) according to the equation

$$CaF_2 + 2\ HCl \text{ -- } CaCl_2 + 2HF$$

If 39 grams of CaF_2 and 73 grams of HCl are mixed together until no further reaction occurs, calculate the mass of hydrogen fluoride (HF) produced." To answer this question the subject must determine the molar mass of each substance from its formula; convert the mass of each reactant to the equivalent number of moles; recognize that one of the reactants is a limiting reagent; use the equation to determine the mole ratio of limiting reagent to desired product; correctly combine this ratio with given mass of reactant to determine number of moles of desired product; and convert this number of moles to the equivalent mass. Such a procedure requires the subject to be able to separate the problem into its component parts, to combine them appropriately, and to use a ratio strategy. Second order operations are involved. Certainly,

viewed from a Piagetian perspective formal thought appears to be necessary for correct solution of the problem. To a slightly lesser extent solution of problems representing Skills 7 and 5, the next lowest skills in the validated hierarchy require similar reasoning abilities. To determine whether individuals must be formal operational to exhibit those skills the Dayton and Macready method was used to contrast the fit of several alternative models to the data.

Each model except Model "a" incorporates developmental level as a prerequisite to one or more intellectual skills. In Model "a" these intellectual skills, which are the uppermost skills in the hierarchy, are independent of developmental level. Because each model has the same components in a different arrangement it is possible, through application of the Dayton and Macready procedure, to compare their goodness of fit directly to the same data. Hence it is possible to compare Model "a" to each of the others and thus to determine whether formal thought is a necessary prerequisite to acquisition of the uppermost skills in the stoichiometry hierarchy. The models and the results of this analysis are presented in Table 19.

Examination of Table 19 indicates that only Model "a" represents a good fit to the data. Further, the difference between the likelihood functions for Model "a" versus Models "b", "c" and "d" is 24.54 (df3), 14.22 (df2) and 15.01 (df1) respectively. This allows acceptance of Model "a" in favour of each of the other models at .00 level of significance. Hence, the data support the hypothesis that formal thought is not necessary to the successful performance of stoichiometric calculations. This is in sharp contrast to theoretical predictions from the perspective of the Piagetian model and also to the finding in chapter

Table 19

Developmental Analysis for Stoichiometry Hierarchy

Model	Maximum Likelihood	Degrees of Freedom	Significance
(a) 9 → 7 → 5 → F	2.78	6	.83
(b) 9 → 7 → 5 → F	27.32	9	.01
(c) 9 → 7 → 5, 7 → F	17.00	7	.01
(d) 9 → 7, 9 → F, 7 → 5	17.79	8	.03

For each model, numbers represent skills and F represents formal thought

four that subordinate intellectual skills, derived by application of a learning hierarchy model, are necessary to acquisition of these same intellectual skills.

The Molarity Hierarchy

Similar procedures were followed for the molarity hierarchy as for the stoichiometry hierarchy. The same developmental tasks were administered in a written format to all subjects in intact classes a few days before testing subjects' competence on the skills in the molarity hierarchy.

The terminal skill in the hierarchy, Skill H, involves appropriate combination of a number of steps. It requires determination of the molarity of a specified common ion when two solutions of specified molarity and volume are combined. An example of this is "2.0L of 0.30MKI, potassium iodide, is added to 3.0L of 0.10 M MgI_2, magnesium iodide. What is the molarity of the iodide ion, I, in the resulting solution?" Solution of this question typically requires the following components: convert molarity of potassium iodide to number of moles of potassium iodide per liter; combine the result of this with the volume of potassium iodide solution to determine the total number of moles of potassium iodide present; combine the result of this with the formula of the potassium iodide to determine the number of moles of iodide ion present in the potassium iodide solution; repeat these three steps for the magnesium iodide solution; add the number of moles of iodide ion present in potassium iodide and magnesium iodide solutions; and combine the result of this with the total volume of solution after mixing to determine the ratio of moles of iodide ion to liters of solution. Solution of this problem requires several applications of a ratio strategy, inferential reasoning,

appropriate combination of a sequence of steps, and application of second-order operations. Likelihood functions, expressed as chi squared values are 2.37 (df=6), 45.78 (df=8) and 45.29 (df7) for Models a, b and c respectively. Again, there is little doubt that, within the Piagetian model, formal thought is required for solution of such problems. To a lesser extent, the same is true for Skill E, which is directly subordinate to Skill H in the validated molarity hierarchy. To test whether formal thought is in fact necessary the Dayton and Macready procedure was applied to three alternative models, which are represented in Table 20. Models "b" and "c" include formal thought as prerequisite to Skills H and F, and Skill H, respectively. In Model "a" developmental level is not prerequisite to either skill. Skill B is retained in each model as part of the overall hierarchical relationship, but is not posited to represent a sufficiently strong formal operational requirement to suggest that formal thought is prerequisite to it. The results of this analysis, represented in Table 20, show that Models "b" and "c" are rejected ($p < .00$) while Model "a" is consistent with the data ($p < .05$). Further, Model "a" is significantly better ($p < .00$) than Model "b" with a chi squared difference of 43.41 (df=2). Similarly, with a difference between chi squared values of 42.89 (df=1), Model "a" is significantly better than Model "c" ($p < .00$). Hence, as in the case of the stoichiometry hierarchy, the ability to exhibit generalized formal-operational thought is not an important determinant of success on items relating to molarity which, according to the Piagetian model, should require such thought processes. As for the stoichiometry hierarchy, this is in sharp contrast to the need to be able to exhibit subordinate skills in order to demonstrate mastery of superordinate skills.

Table 20

Developmental Analysis for Molarity Hierarchy

Model	Maximum Likelihood	Degrees of Freedom	Significance
(a) H → E → B → F	2.37	6	>.95
(b) H → E; E → B; E → F	45.78	8	<.00
(c) H → E; H → F; E → B	45.29	7	<.00

For each model H, E and B represent skills and F represents formal thought

The Food Web Hierarchy

Similar procedures to those used for the stoichiometry and molarity hierarchies were again applied. The terminal skill in the hierarchy, Skill 9, requires the subject, given a food web diagram, to "determine the effect of a sudden size change in one population on a second population which is not on the same food chain, when the effect may be transmitted along more than one route." Given that the two populations are not part of a common food chain, in order to exhibit this skill correctly the subject must realize that all populations in a food web affect each other regardless of their location in the web. The particular test questions used were limited to webs involving no more than four possible routes between the populations in question. Even so, the subject must in effect make a number of qualified inferences about each particular possible route between these populations, and combine these inferences to produce an overall answer. Skills 7 and 8 require similar reasoning but within one food chain. For all three skills the subject must combine several subordinate rules, apply combinatorial reasoning to all possible combinations of pathways, apply hypothetico-deductive logic to arguments about a series of connections, and be able to reflect on his own reasoning. As for the stoichiometry and molarity concepts, within the Piagetian model formal thought appears to be necessary to acquisition of the upper skills in the hierarchy. Whether Skill 6, "given a food web diagram indicate all possible pathways through which the effect of a change in one population is transmitted to a second population", involves formal thought is more open to question. However, it does involve simultaneous consideration of a number of possibilities and was included in a superordinate position relative to formal

thought in one of the alternative models tested. To determine whether formal thought is necessary to acquisition of the four top skills in the hierarchy each model was compared to Model "a", which does not include developmental level. The results of these analyses are represented in Table 21.

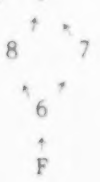
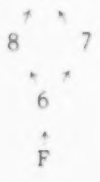
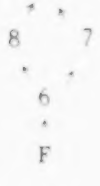
There were no exceptions to Model "a" in which formal thought was hypothesized not to be required for acquisition of any of skills 6, 7, 8 or 9. Hence the chi square value for goodness of fit between this model and the data was zero (df = 16). In contrast, for Model "b" in which formal thought was hypothesized to be prerequisite to four skills, 38 percent of the response patterns were anomalous. The chi square for this model was 104 (df = 23). The difference between the chi square values for Models "a" and "b" is clearly also significant (X^2 = 104, df = 7), and it can be concluded that Model "a" is superior to Model "b" ($p < .01$).

Model "c", in which formal thought is hypothesized to be prerequisite to all skills except Skill 6 has 26% anomalous response patterns and a chi square of 93.5 (df = 22). With a difference between chi square values of 93.5 (df = 6), Model "c" is inferior to Model "a" ($p < .00$). Finally, for Model "d", in which formal thought is suggested to be prerequisite only to the terminal skill, 12 percent of the response patterns were anomalous. Again the model does not offer a good fit to the data (X^2 = 58.4, df = 22), and the difference in chi square values (58.4, df = 6) between this model and Model a is significant at the .00 level.

In summary, as for the stoichiometry and molarity hierarchies, the ability to exhibit formal thought is not, contrary to what would be

Table 21

Developmental Analysis for Food Web Hierarchy

Model	Maximum Likelihood	Degrees of Freedom	Significance
(a) 	0.00	16	1.00
(b) 	104.0	23	<.00
(c) 	93.5	22	<.00
(d)	58.4	22	<.00

For each model, numbers represent skills and F represents formal thought.

inferred from Piagetian theory, prerequisite to the ability to demonstrate the uppermost skills in the hierarchy. Again, this is in marked contrast to the need for prior learning of subordinate intellectual skills.

The Conservation of Mechanical Energy and Mendel's Laws Hierarchies

Data are not presented for the dependence of these skills on formal thought. In the case of the sample for the mechanical energy hierarchy too few subjects (less than ten percent of the sample) were deemed non-formal to allow for meaningful analysis. For the Mendel's Laws hierarchy the opposite problem emerged. Too few subjects, whether formal or not, were able to exhibit the upper skills to allow for meaningful analysis.

Summary

In the present chapter data have been presented which show that for three of the five concepts investigated, namely the concepts of stoichiometry, molarity and food web, subjects' ability to exhibit formal thought was not prerequisite to their ability to demonstrate intellectual skills which, from a Piagetian analysis, demand formal thought. The form of the data for two more concepts, conservation of mechanical energy and Mendel's Laws, respectively, did not allow for similar analysis for these concepts.

Chapter Six

Epilogue and Prologue

The study which has been reported offers a link between the past and the future. It derives from Gagne's introduction of the concept of learning hierarchies in the sixties and from his subsequent restriction of this to the domain of intellectual skills. It derives also from the work of Piaget, in particular his contention that learning is most dependent upon the learner's stage of intellectual development. In the present study, for the narrow range of learning which Gagne claimes to be best represented in learning hierarchies and, further, primarily with respect to the adolescent learner, these two theoretical positions were contrasted. The results supported the learning hierarchy model over the developmental model. Learning was found to be dependent upon specific prerequisites but not upon the availability of formal thought, even when a Piagetian analysis suggested that extensive formal thought should be necessary. It is important to take this finding for just what it is. It emphasizes the importance of related prior learning to the learning of what are essentially type problems. Further, it illustrates the relative lack of importance of the ability of the learner to display formal thought in situations where the learner has been exposed to and mastered the necessary prerequisites. Other kinds of learning such as novel problem solving might yield a quite different result, and it is clearly not possible to generalize beyond the kind of content nor perhaps beyond the adolescent learner as studied in the present study.

Even a cursory examination of science courses at the high school level and beyond illustrates the importance of type problems to basic learning of many science concepts. Such problems are particularly

amenable to hierarchial analysis. The present study is also important, then, in adding to the literature a validated hierarchy for each of five important science concepts.

In addition to the findings described in this report, the wealth of student responses to test items representing the intellectual skills studied yielded a direction for future related research. Careful examination of student's responses indicated a number of consistent errors which could be interpreted in terms of underlying misconceptions. Combination of these together with the hierarchies validated in the present study formed the basis for a further study which has already been funded and substantially completed, in which students individual misconceptions for the concepts studied in the present study are exposed as a means to remediation. A report on this further study will be available soon. Such a study is consistent with the universal interest being shown at present in students' misconceptions of science concepts. Finally, the reader who is interested in further detail of the research describe in the present study and the concurrent identification of learner misconceptions is directed to four M.Ed. theses supervised by the author. These are, for the stoichiometry concept Whelan (1982), for the molarity concept Hopkins (1985), for the conservation of mechanical energy concept Pottle (1982), for the food web concept Grant (1985). No thesis was submitted for the Mendel's laws hierarchy. None of the theses contains information about the Piagetian aspects of the study. Copies of all tests and remedial units used in the overall study are available from the principal investigator. The relationship between the present study and each of the thesis studies was truly symbiotic. I wish to thank again each of these researchers for the learning they brought to me.

References

Airasian, P.W., & Bart, W.M. (1975). Validating a priori instructional hierarchies. Journal of Educational Measurement, 12(3), 163-173.

Bart, W.M. & Krus, D.J. (1973) An ordering-theoretic method to determine hierarchies among items. Educational and Psychological Measurement, 33, 291-300.

Beeson, E.W.(1977) Hierarchical learning in electrical science. Journal of Research in Science Teaching, 14(2), 117-127.

Beilin, H. (1971) The training and acquisition of logical operations. In Piagetian cognitive development and mathematics education, M.F. Rosskopf (Ed.), National Council of Teachers of Mathematics, Washington, D.C.

Bergan, J.R. (1980). The structural analysis of behavior: An alternative to the learning-hierarchy model. Review of Educational Research. 50(4), 625-646.

Boyle, R.P. (1970). Part analysis and ordinal data. American Journal of Sociology 75(4,1) 461-480.

Brainerd, C.J. (1973). Neo-Piagetian training experiements revisited: Is there any support for the cognitive-developmental stage hypothesis. Cognition, 2(3), 349-370.

Briggs, L.J. (1968). Sequencing of instruction in relation to hierarchies of competence. Pittsburgh: American Institute for Research.

Capie, W. & Jones, H.W. (1971). An assessment of hierarchy validation techniques. Journal of Research in Science Teaching, 8, 137-147.

Case, R. (1974). Structures and strictures, some functional limitations on the course of cognitive growth. Cognitive Psychology, 6, 544-573.

Case, R. (1978). A developmentally based theory and technology of instruction. Review of Educational Research, 48, 439-469.

Cotton, J.W., Gallagher, J.P., and Marshall S.P. (1977) The identification and decomposition of hierarchical tasks. American Educational Research Journal, 14(93), 189-212.

Dayton, C.M. & Macready, G.B. (1976a). A probabilistic model for validation of behavioral hierarchies. Psychometrika, 41(2), 189-204.

Dayton, C.M. & Macready, G.B. (1976b). Computer programs for probabilistic models. University of Maryland.

Eisenberg, T.A. & Walbesser, H. (1971). Learning Hierarchies - numerical considerations. Journal for Research in Mathematics Education, 2(4), 244-256.

Gagne, R.M. (1968). Contributions of learning to human development. Psychological Review, 75(3), 177-191.

Gagne, R.M. (1970a). Some new views of learning and instruction. Phi Delta Kappan, 51(9), 468-472.

Gagne, R.M. (1970b). The conditions of Learning (2nd ed.). New York: Holt, Rinehart and Winston.

Gagne, R.M. (1972). Domains of learning. Interchange, 3, 1-8.

Gagne, R.M. (1973). Learning and instructional sequence. In F. Kerlinger (Ed.), Review of Research in Education, (1), Itasca, Illinois: F.E. Peacock.

Gagne, R.M. (1977). The conditions of Learning (3rd ed.). New York: Holt, Rinehart and Winston.

Gagne, R.M. & Briggs, L.J. (1974). Principles of instructional design, New York: Holt, Rinehart & Winston.

Gagne, R.M. & Paradise, M.E. (1961). Abilities and learning sets in knowledge acquisition. Psychological Monographs, 75 (whole No. 518).

Goldschmid, M.L. (1967). Different types of conservation and non-conservation and their relation to age, sex, IQ, MA, and vocabulary. Child Development, 38 (1229-1246).

Grant, B.A.C. (1983). The food web concept: Identification of a learning hierarchy and related misconceptions. Unpublished master's thesis, Memorial University of Newfoundland, St. John's.

Griffiths, A.K. (1979). The mole concept: Investigation of an hierarchical model. Unpublished doctoral dissertation, University of Alberta.

Griffiths, A.K. and Grant, B.A.C. (1985). High school students understanding of food webs: identification of a learning hierarchy and related misconceptions. Journal of Research in Science Teaching, 22, 421-436.

Griffiths, A.K., Kass, H., & Cornish, A.G. (1983). Validation of a learning hierarchy for the mole concept. Journal of Research in Science Teaching, 20, 639-654.

Griffiths, A.K., Pottle, J.E., and Whelan, P.H. (1983). Application of the learning hierarchy model to the identification of specific misconceptions for two science concepts. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Dallas.

Hobbs, E.T. (1975). Formal operations in secondary students. Unpublished doctoral dissertation, University of Alberta.

Hopkins, B.L. (1985). The molarity concept: identification of a learning hierarchy and related students errors. Unpublished M.Ed. dissertation, Memorial University of Newfoundland.

Inhelder, B. & Piaget, J. (1958). The growth of logical thinking. New York: Basic Books.

Jones, H.J. & Russell, J.M. (1979). Hierarchical learning paradigm. Journal of Research in Science Teaching. 16 (6), 489-499.

Johnstone, A.H. (1979). The content-a barrier to implementation, occasional paper, University of Glasgow.

Karplus, Robert. (1977). Science teaching and the development of reasoning. Journal of Research in Science Teaching. 14(2), 169-175.

Karplus, E.F. & Karplus, R. (1970) Intellectual development beyond elementary school. School Science and Mathematics, 70, 398-406.

Karplus, R & Lavatelli, C. (1969). The developmental theory of Jean Piaget. San Francisco: Davidson films.

Lawson, A.E. (1978). The development and validation of a classroom test of formal reasoning. Journal of Research in Science Teaching, 15(1), 11-24.

Lovell, K. (1980). The relevance of cognitive psychology to science and mathematics education, in Cognitive Development Research in Science and Mathematics, W.F. Archenhold (Ed.), Leeds: University of Leeds.

Linke, R.D. (1975), Replicative studies in hierarchical learning of graphical interpretation skills. British Journal of Educational Psychology, 45, 39-46.

Nagy, P. and Griffiths, A.K.(1982). Limitations of recent research relating to Piaget's theory of adolescent thought, Review of Educational Research, 52(4), 513-556.

Novak, J.D. (1977). A Theory of education. Ithaca, New York: Cornell University Press.

Okey, J.R. & Gagne, R.M. (1970). Revision of a science topic using evidence of performance on subordinate skills. Journal of Research in Science Teaching, 7, 321-325.

Piaget, J. & Inhelder, B. (1962). Le developpment des quantities physiques chez l'enfant (2nd ed.). Neuchatel: Delachause et Neustel.

Owston, R.D. (1979). A maximum likelihood approach to the test of inclusion. Psychometrika, 44(4), 421-425.

Piaget, J. & Inhelder, B (1975). The origin of the idea of chance in children. London: Routledge & Keagan Paul.

Posner, G.J. & Strike, K.A. (1976). A categorization scheme for principles of sequencing content. Review of Educational Research, 46(4), 665-690.

Pottle, J.E. (1982). A study investigating a hierarchial model relating to the concept of conservation of mechanical energy. Unpublished master's thesis, Memorial University of Newfoundland, St. John's.

Renner, J.W. Evaluating intellectual development using written responses to selected science problems. Norman, Oklahoma: University of Oklahoma, 1977.

Renner, J.W. & Stafford, D.G. (1972). Teaching science in the secondary school. New York: Harper and Row.

Resnick, L.B. & Wang, M.C. (1969), Approaches to the validation of Learning hierarchies. Educational Testing Service.

Shayer, M., Adey,P. & Wylam, H. (1981) Group tests of cognitive development: ideals and a realization. Journal of Research in Science Teaching. 18(2), 157-168.

Strauss, S. (1972). Learning theories of Gagne and Piaget: Implications for curriculum development. Teachers College Record, 74(1), 81-102.

Strike, K.A. & Posner, G.J. (1976). Epistemological perpectives on conceptions of curriculum orgization and learning, in Review of Research in Education, L.S. Shulman (Ed.). Itasca, Illinois: F.E. Peacock Inc.

Walbesser, H.H. & Eisenberg, T.A. (1972). A review of research on behavioral objectives and learning hierarchies. (ERIC Document 059 900).

Whelan, P.J. (1982). A study investigating a hierarchial model relating to stoichiometric calculations. Unpublished master's thesis, Memorial University of Newfoundland, St. John's.

Wheeler, A.E. (1976) Proportionality in high school chemistry. Unpublished doctoral dissertation. University of Alberta.

White, R.T. (1973). Research into learning hierarchies. Review of Educational Research. 43(3), 361-375.

White, R.T. (1974a) The validation of a learning hierarchy. American Educational Research Journal. 11, 121-136.

White, R.T. (1974b) A model for validation of learning hierarchies, Journal of Research in Science Teaching. 11(1), 1-3.

White, R.T.(1974c) Indexes used in testing the validity of learning hierarchies. Journal of Research in Science Teaching. 11(1), 61-66.

White, R.T. & Clark, R.C. (1973). A test of inclusion which allows errors of measurement. Psychometrika, 38, 77-86.

Wollman, W. (1977) Controlling variables: assessing levels of understanding. Science Education. 61(3), 385-391.